Kernel Correlation Feature Analysis: A New Advanced Correlation Filter Approach for Recognizing Uncontrolled Face Image data in FRGC - Phase II

Dr. Marios Savvides

Email: Marios.Savvides@ri.cmu.edu
www.cs.cmu.edu/~msavvid

Research Assistant Professor
Electrical & Computer Eng and CyLab
Carnegie Mellon University

Acknowledgements

- In collaboration with Prof. VijayaKumar at CMU (we are participating in FRGC and FRVT!).

- Thanks to Dr. Jonathon Phillips and NIST for creating these interesting face challenge problems and pushing to create the next generation face recognition algorithms!

- Very grateful for support of this research by the U.S. Technical Support Working Group (TSWG)!

Challenges in Face Recognition and how related to Quality?

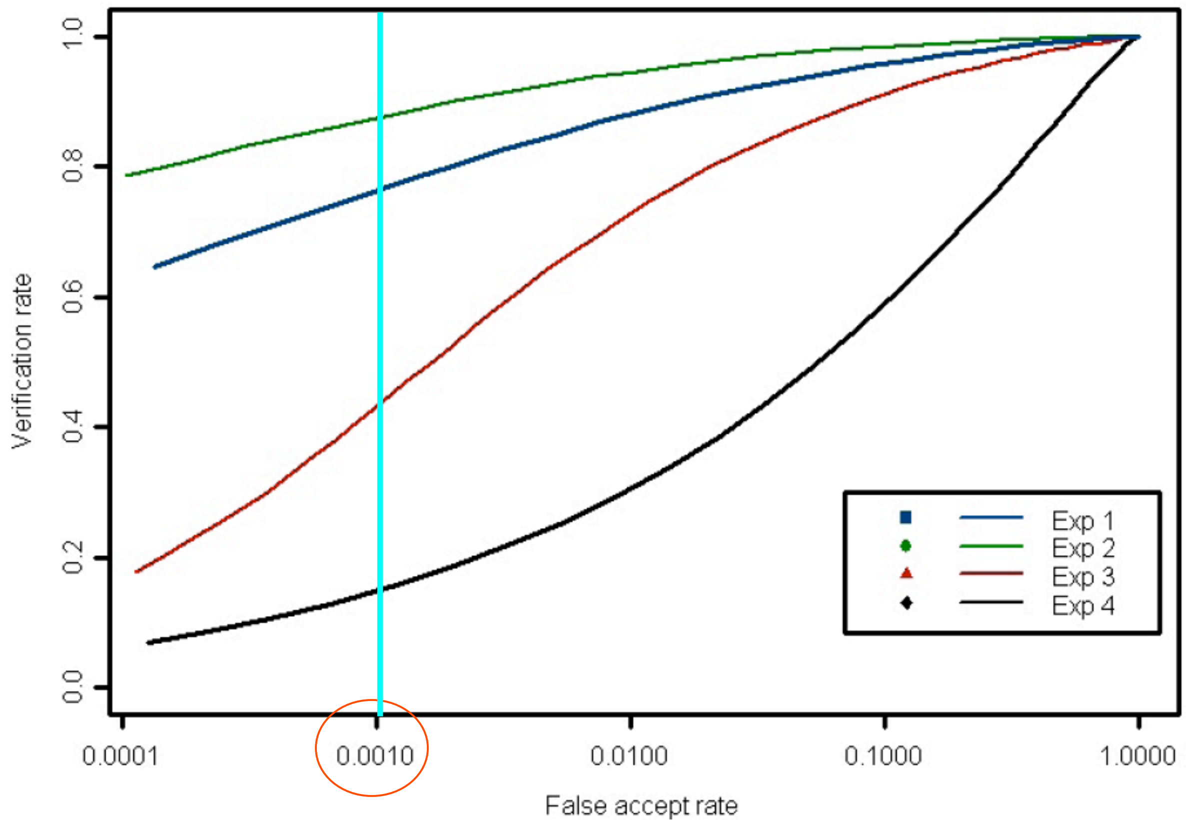
- Pose
- Illumination
- Expression
- Occlusion
- Time lapse
- Individual factors: Gender

Motivation

- Face recognition is a challenging task due to large distortions in human faces.
- Face Recognition Grand Challenge (FRGC) program evaluates the state of the art of current face recognition technology.
- Lots and Lots of face data captured over a period of 2 years at University of Notre Dame (>36,000 images).
- Participation includes commercial face recognition vendors and Universities.
- We want to contribute our efforts to develop more robust face recognition algorithms with large scale database such as the FRGC.
- We have to design new approaches to recognizing faces with correlation filters
- Current Correlation Filter methods do not use generic training set, we present a novel approach that uses generic set to build a generic CF basis

FRGC Experiments

Principal Component Analysis (PCA) Performance



The performance of PCA is about 12% at the False Accept Rate 0.1%.

➢ The Experiment 4 is most challenging due to uncontrolled conditions

ROC curve from P. Jonathan Phillips et al (CVPR 2005)

FRGC Gallery/Target Images

Controlled (Indoor)

These images are what you have of the 'criminal' suspect that you are looking for...

(we have 16,028 target images of 466 people)

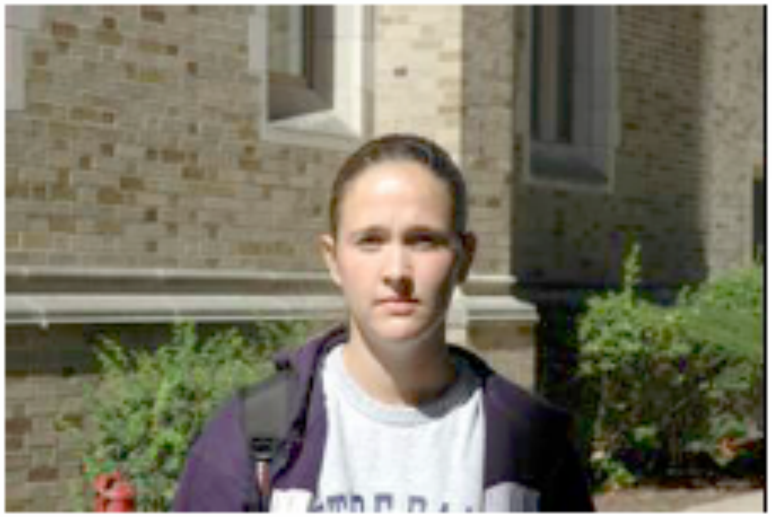
FRGC Query Images

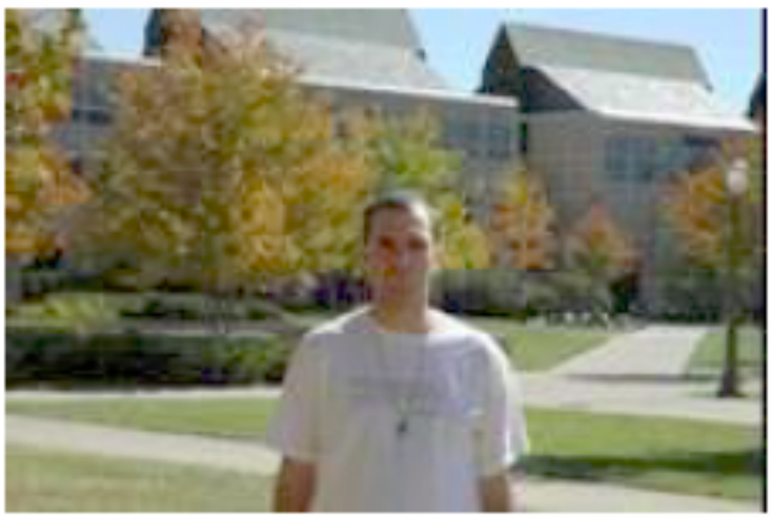
Uncontrolled (Indoor)

These are the test images captured un uncontrolled conditions that we must be able to match against the 'Target' set

more FRGC Query Images

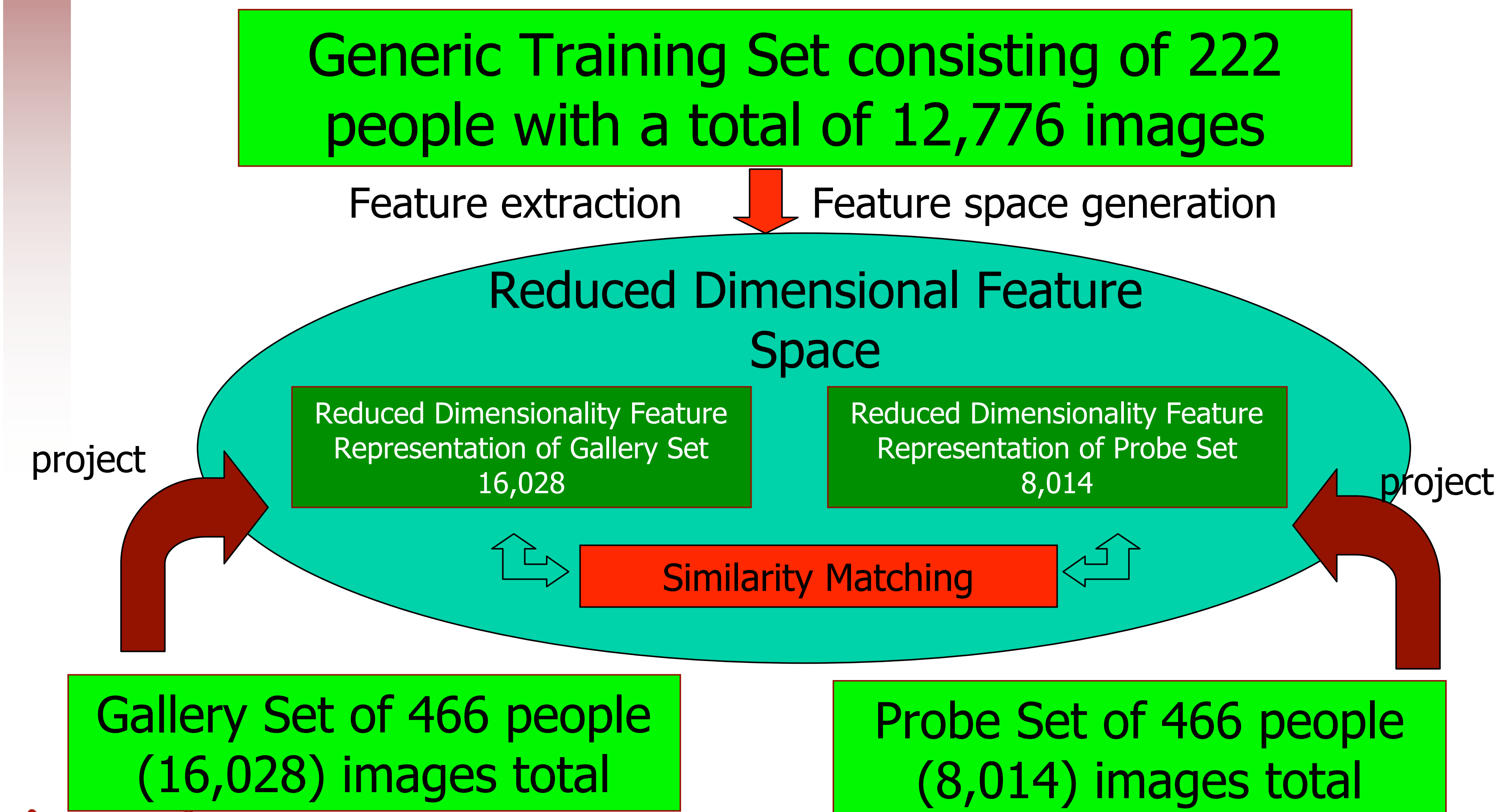
Uncontrolled (Outdoor)





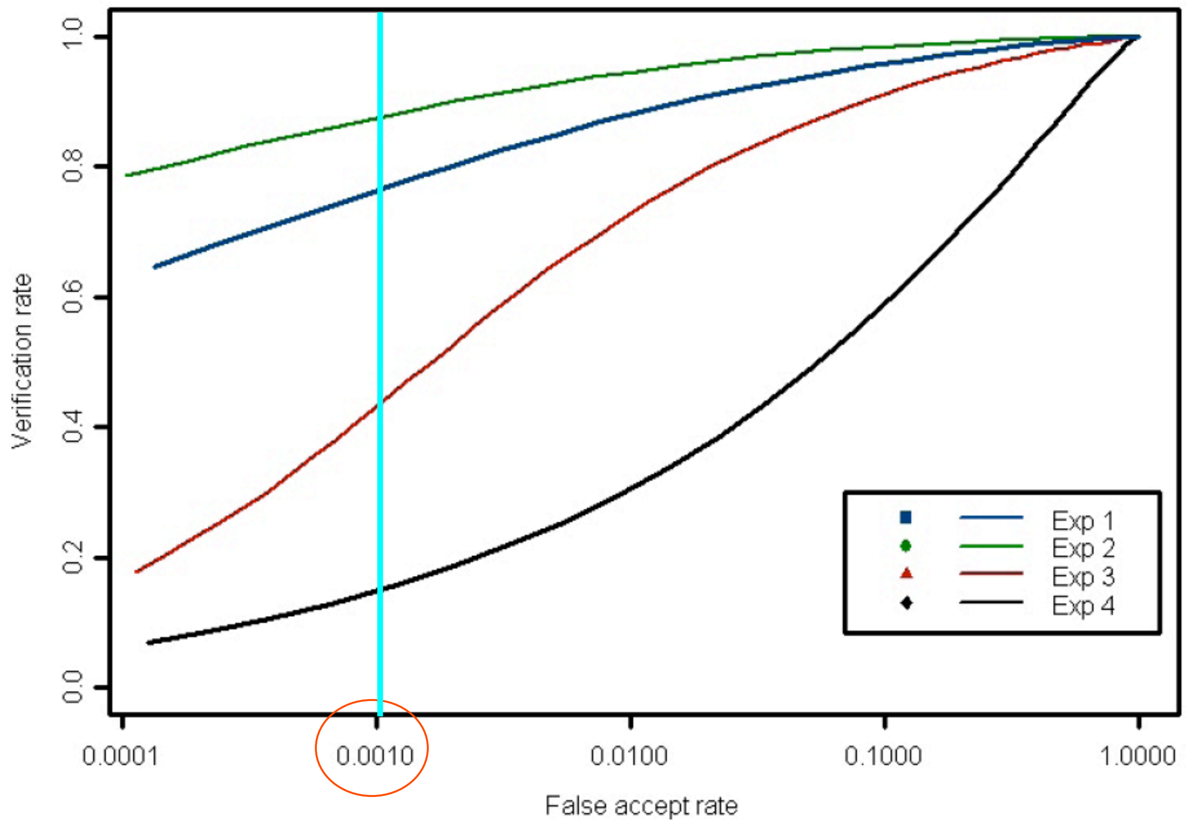
Outdoor illumination images are very challenging due to harsh cast shadows, these affect image quality significantly.

Face Recognition Grand Challenge Dataset from NIST



FRGC Experiments

Principal Component Analysis (PCA) Performance



The performance of PCA is about 12% at the False Accept Rate 0.1%.

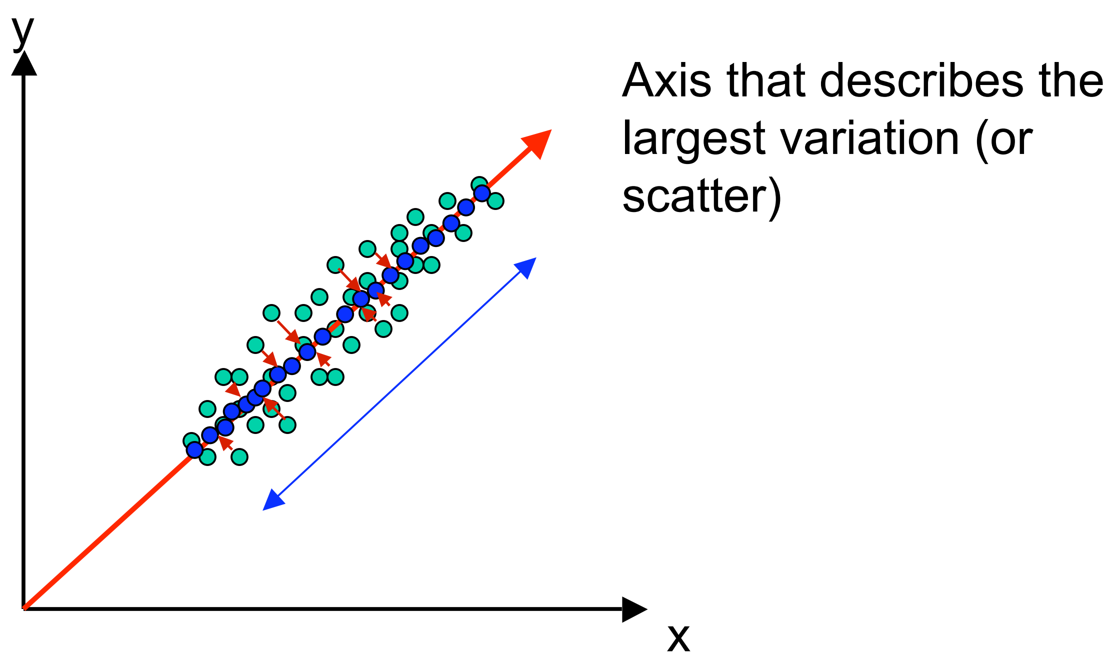
- **The Experiment 4 is most challenging due to uncontrolled conditions. This is what we focus on….**

Electrical & Computer ENGINEERING

ROC curve from P. Jonathan Phillips et al (CVPR 2005)

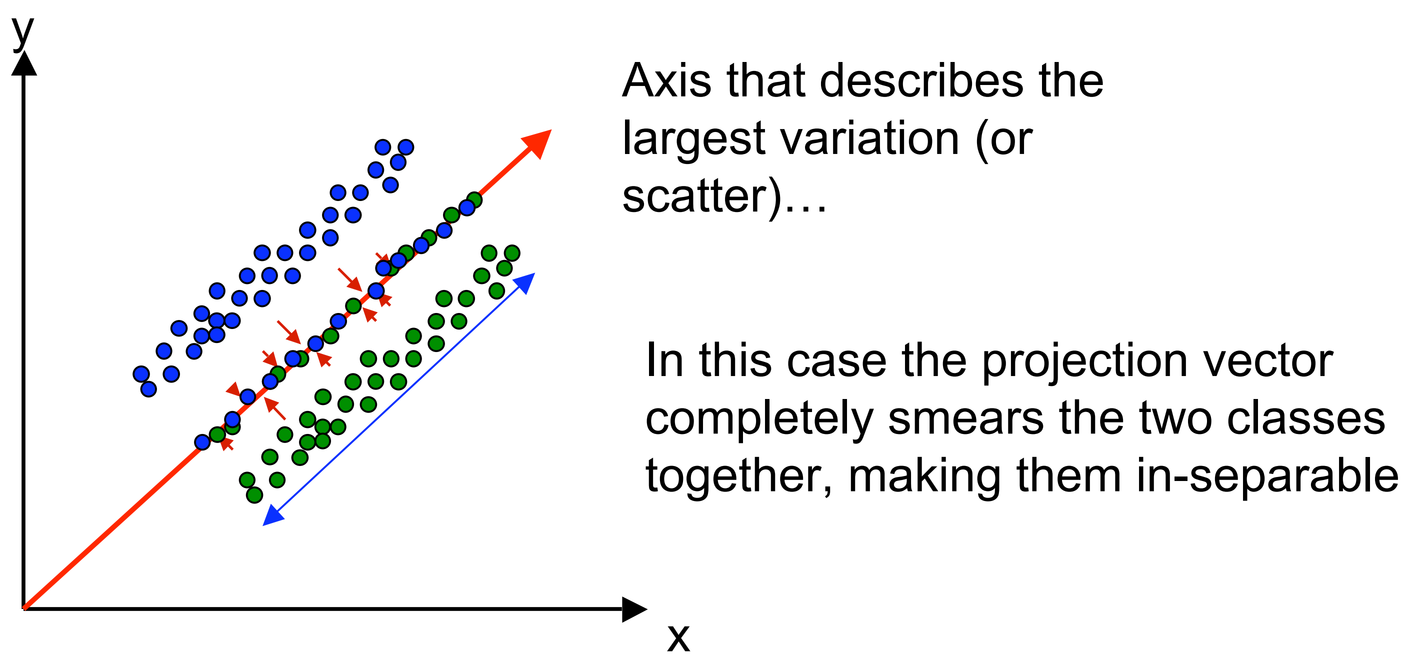
What is PCA? What is it trying to do?

- We want to find projections of data (i.e. direction vectors that we can project the data on to) that describe the maximum variation (or capture most of signal energy).
- Hopefully we can represent the data with a few such projection vectors.
- PCA is very one of the most typically dimensionality reduction methods used in pattern recognition.



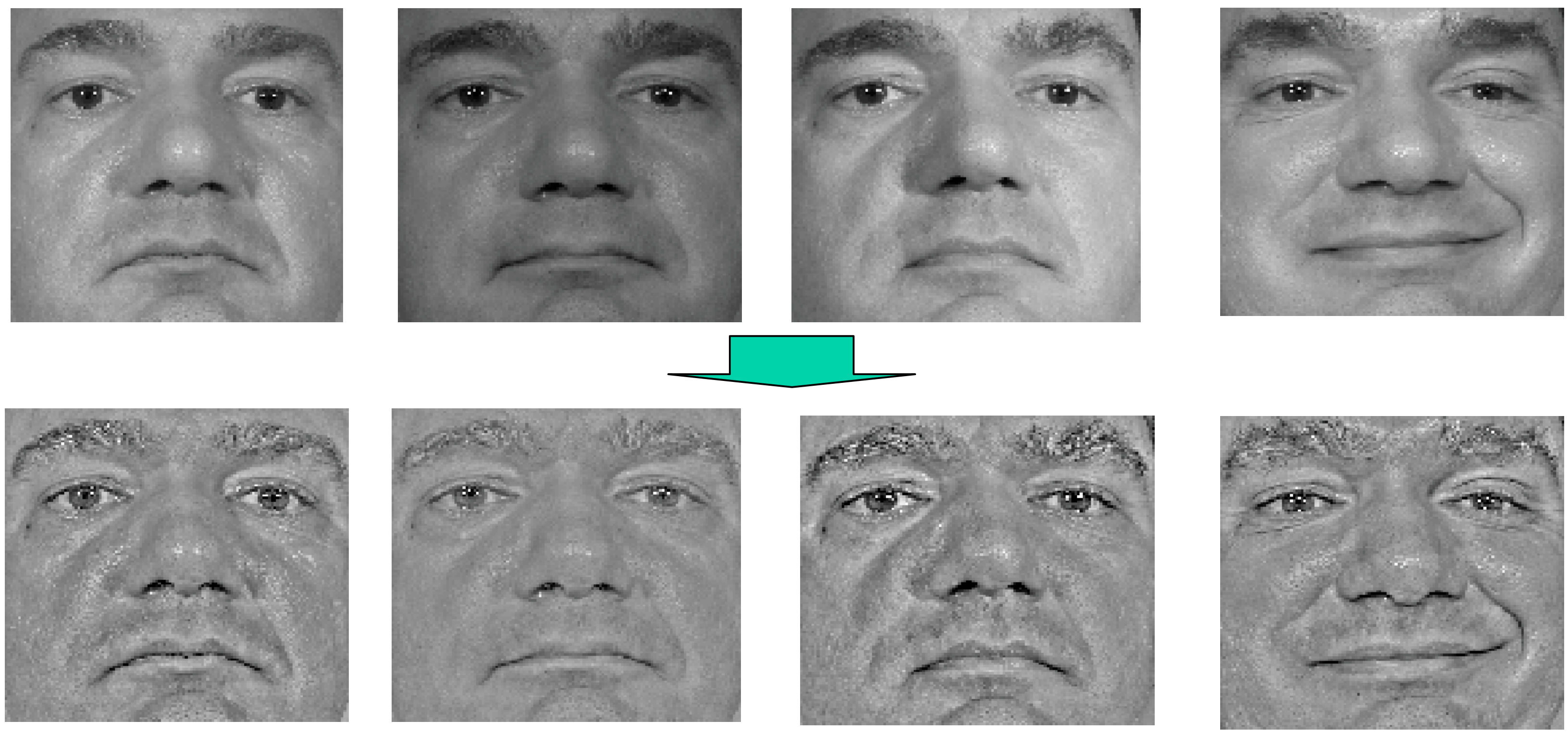
When is PCA bad?

- What if we have 2 classes (the green and blue dots and we want to separate them in some feature space)?
- This is what PCA does..



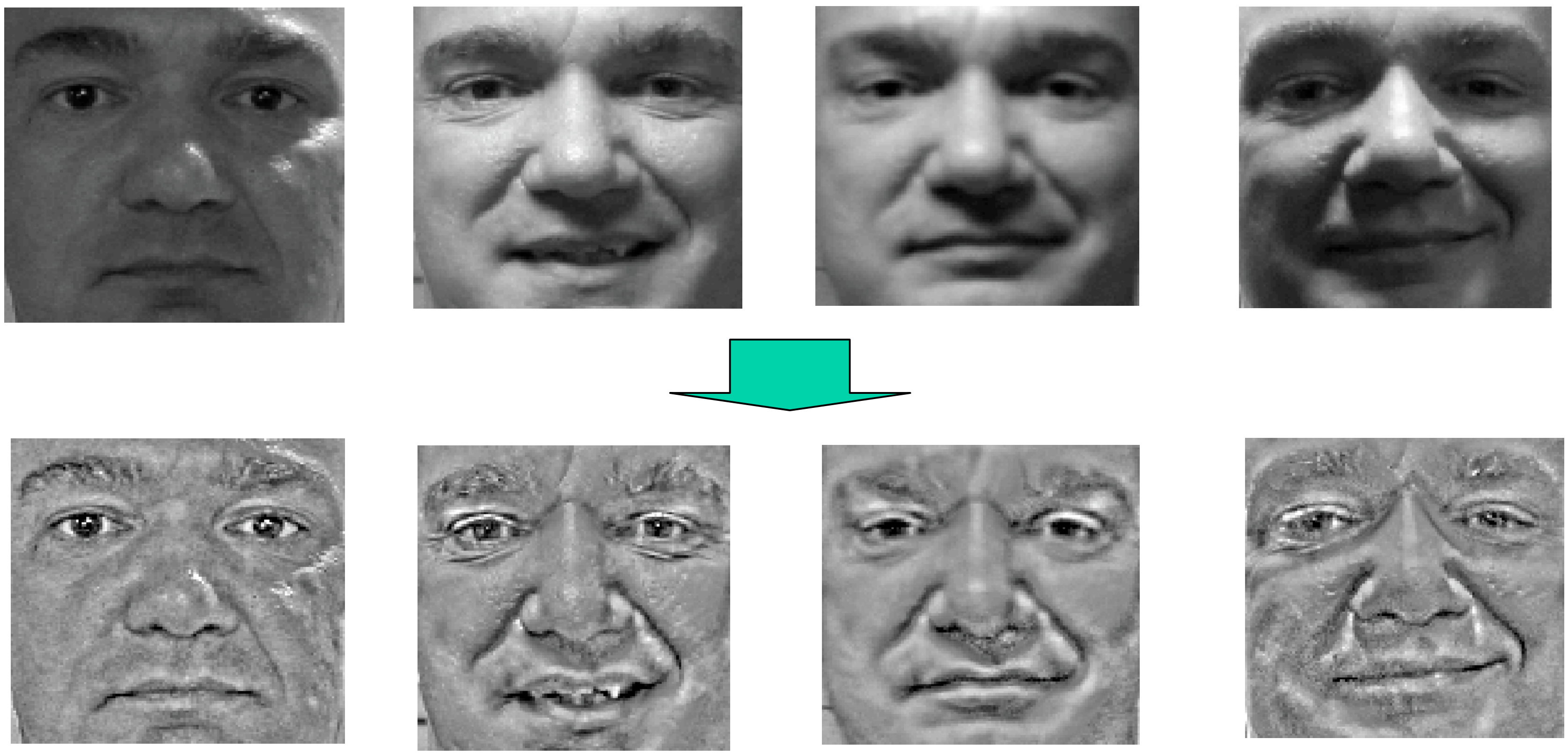
Illumination Pre-Processing

- We examined different illumination algorithms and we used one from CMU (Gross & Brajovic)



Controlled Face Images Illum-processed

Illumination Pre-Processing



Uncontrolled Face Images (harsh illumination conditions) Illum-processed

Why did PCA fail in FRGC ExpIV?

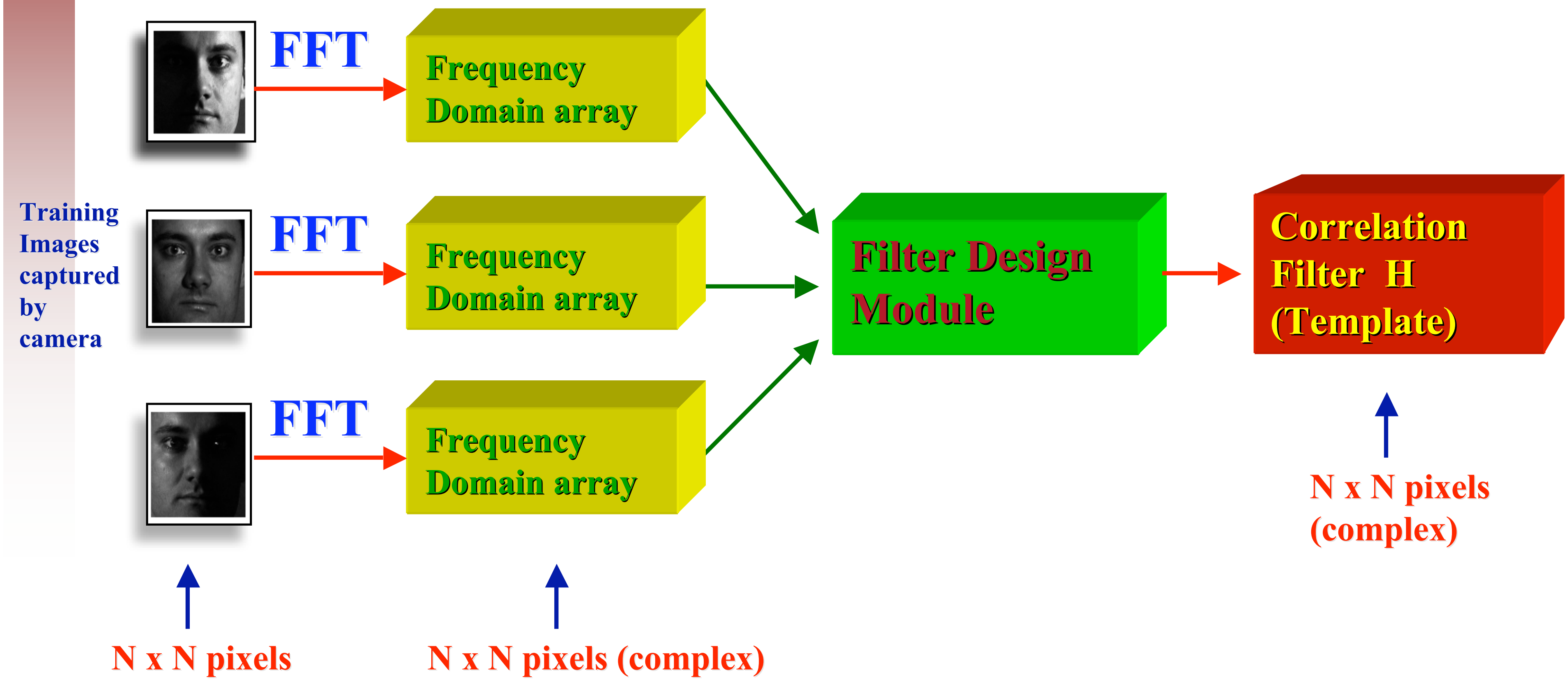
- Even though PCA subspace was built using 12,776 images from 222 people…..
- Final verification rate was very low (12% @ 0.1 FAR).
- This suggests PCA subspace could not represent the gallery/probe images in generated subspace.
- Poor discrimination ability.
- Illumination pre-processing did not seem to help PCA much.

How about Advanced Correlation Filter Designs?

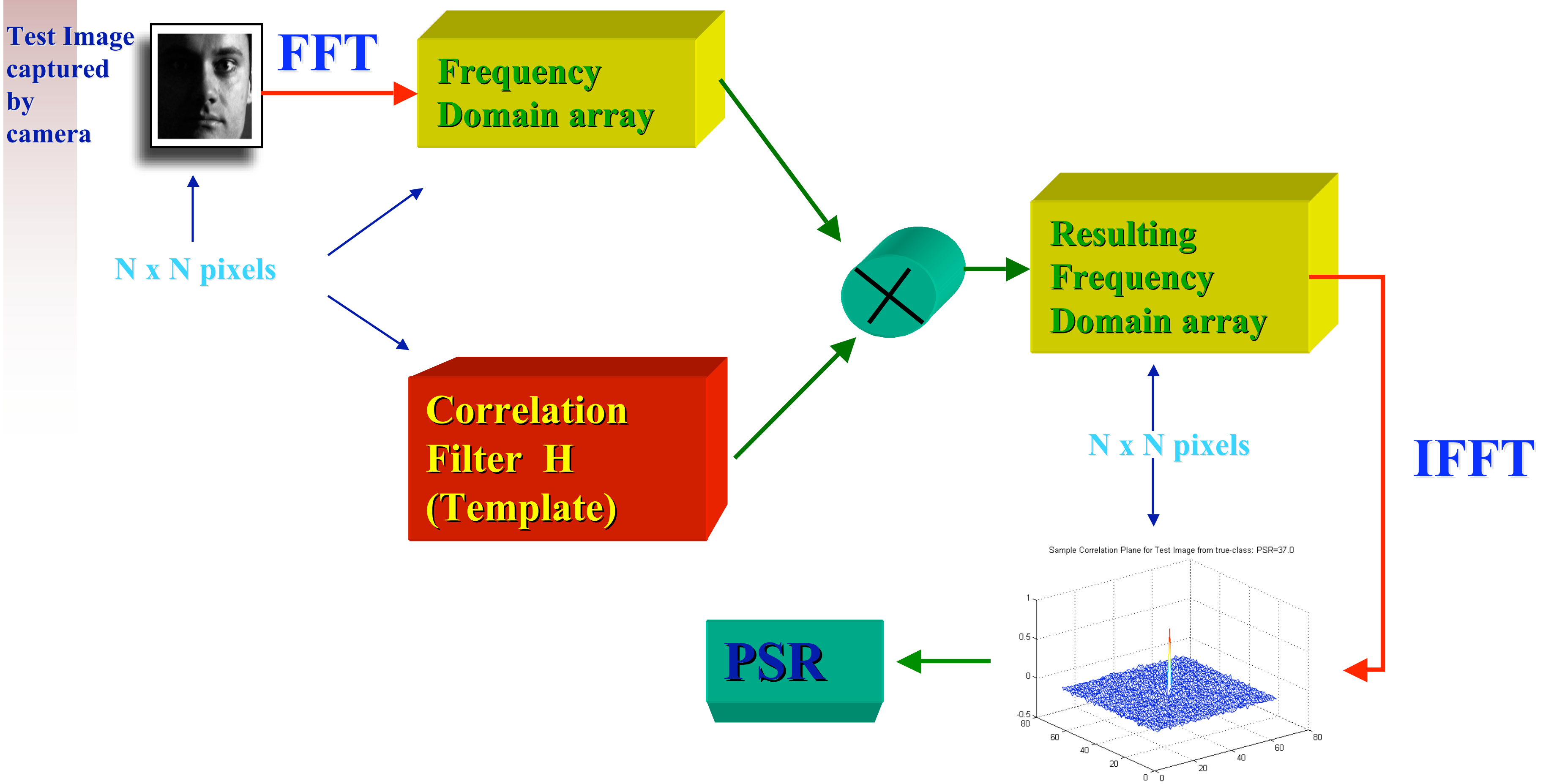
- Advanced Correlation Filter Designs in past such as Minimum Average Correlation Energy (MACE) filter and its derivatives worked well for illumination tolerant face recognition on CMU-Pose Illumination Expression (PIE) database.

- How can they be used successfully in FRGC?

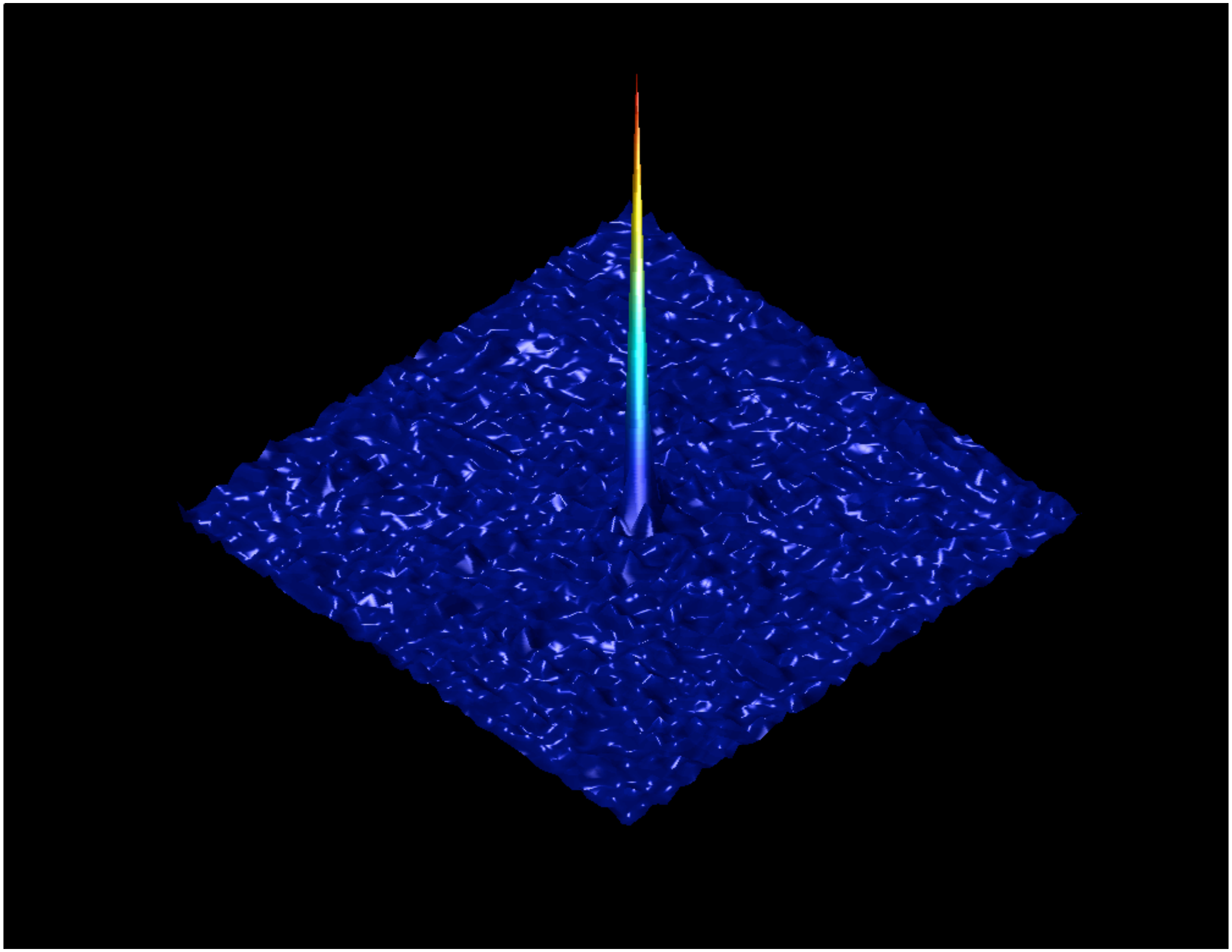
Typical Enrollment Scenario for Correlation Filter Recognition



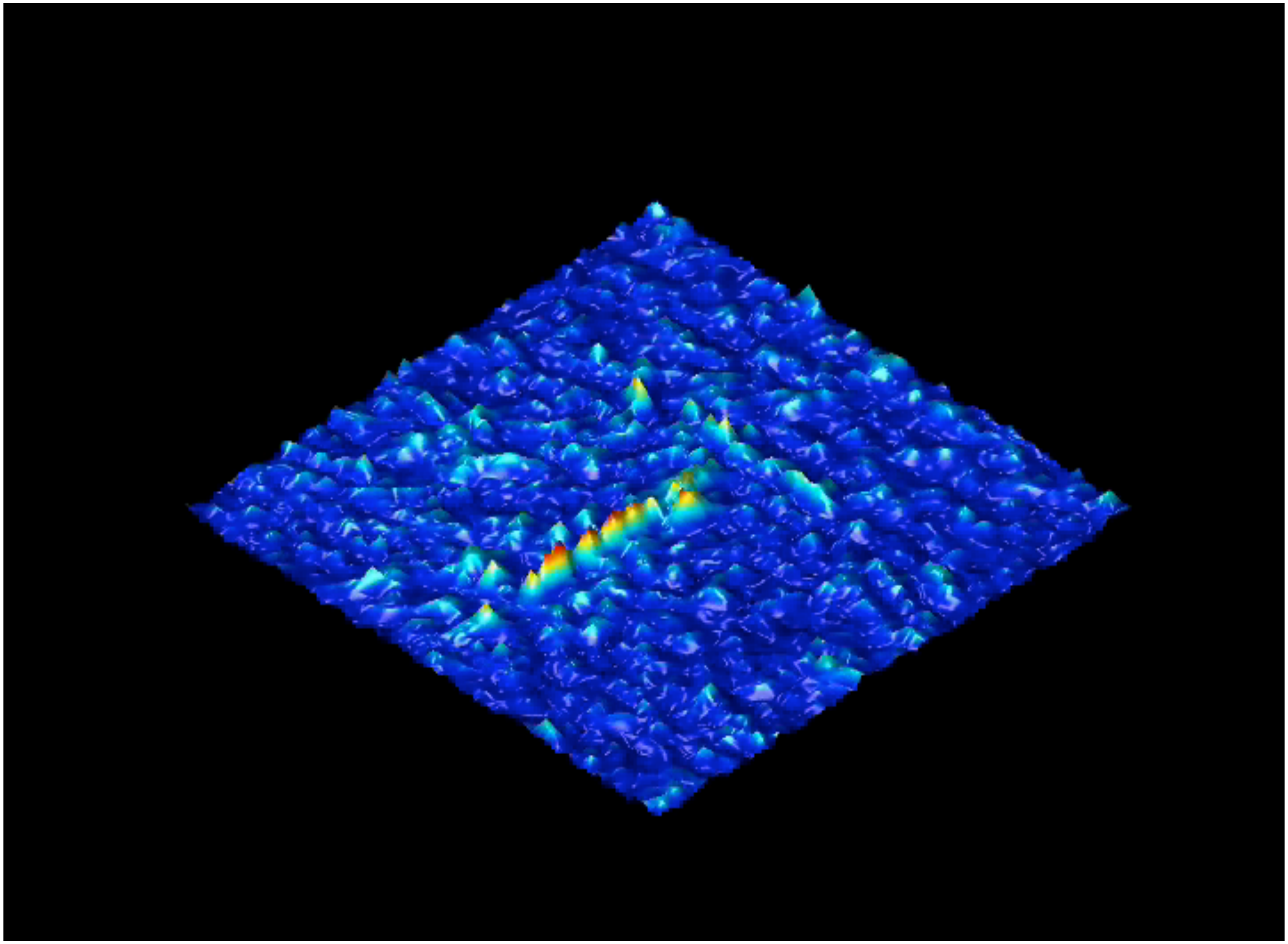
Recognition stage (traditional way of doing correlation)



Typical Correlation Outputs from an Authentic

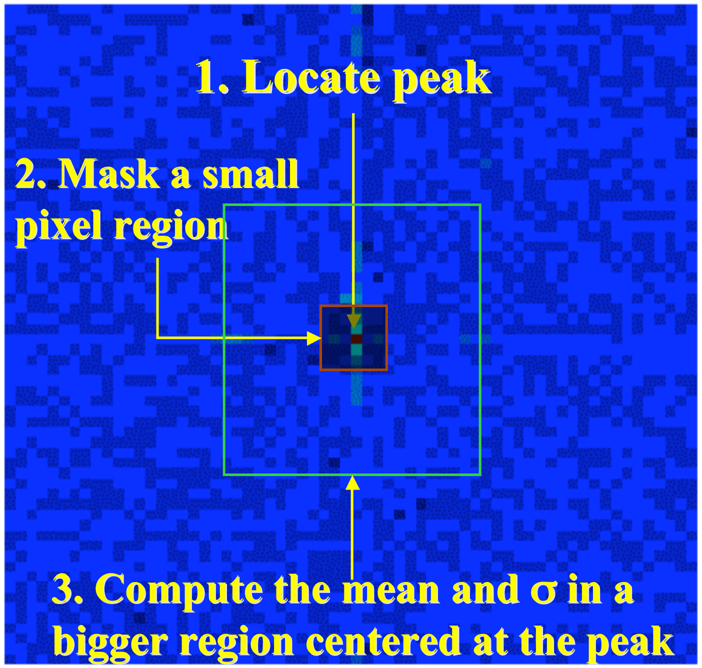


Example Correlation Outputs from an Impostor



Peak to Sidelobe Ratio (PSR)

- PSR invariant to constant illumination changes



$$PSR = \frac{Peak - mean}{\sigma}$$

- Match declared when PSR is large, i.e., peak must not only be large, but sidelobes must be small.

What is the problem so far?

- Typical Correlation Filter Configurations do not make use of "generic" training data.
- There is no notion of "generic filter basis"
- We have a Target, we build a filter to look for it in a cluttered scene…..end of story…………

……or is it?

Different Approach to Achieve Dimensionality reduction

- Instead of finding representation coefficients, let us look at cognitive inspired approach.
- How do we discriminate people (new) people versus people we have seen before.
- One explanation can be we learn **discriminative features between people** we know.
- Once a new person is presented to us we find out how or **who he resembles (via correlations)** based on the discriminative features of the people we know.

From our memory of people we have seen, who does the Query subjects look like?

•Possibly a cognitive inspired approach…

•How do you recognize or relate features of a new person you haven't seen?

Generic Training Images

0.1 0.8 0.05 0.3

Probe Test image

Feature similarity Vector

0.1
0.8
0.05
0.3

Close to 1=>very similar

Close to 0 => little similarity

How do we do this in practice?

- We train advanced correlation filters in a discriminative way to produce an orthogonal feature space.
- Each person has a filter which has been trained to yield +1 correlation output for that person and 0 for all other people.
- Each person's CF learns the discriminative features of that person compared to every other person we have in the database.
- Thus projection features on these CFs gives us a measure of how a new (probe) person resembles that particular person only.

How we design Filter 1 (i.e. for person 1)

$<$  $, \mathbf{h_1} > \; = \; 1$

$<$  $, \mathbf{h_1} > \; = \; 1$

$<$  $, \mathbf{h_1} > \; = \; 1$

$<$  $, \mathbf{h_1} > \; = \; 0$

$<$  $, \mathbf{h_1} > \; = \; 0$

$<$  $, \mathbf{h_1} > \; = \; 0$

$<$  $, \mathbf{h_1} > \; = \; 0$

$<$  $, \mathbf{h_1} > \; = \; 0$

$<$  $, \mathbf{h_1} > \; = \; 0$

How we design Filter 2

<  , h_2 > = 0

<  , h_2 > = 0

<  , h_2 > = 0

<  , h_2 > = 1

<  , h_2 > = 1

<  , h_2 > = 1

<  , h_2 > = 0

<  , h_2 > = 0

<  , h_2 > = 0

How we design Filter 3

$<$  $, h_3 > = 0$

$<$  $, h_3 > = 0$

$<$  $, h_3 > = 0$

$<$  $, h_3 > = 0$

$<$  $, h_3 > = 0$

$<$  $, h_3 > = 0$

$<$  $, h_3 > = 1$

$<$  $, h_3 > = 1$

$<$  $, h_3 > = 1$

How to do feature extraction

Target/Query Image

$\langle$  $, \mathbf{h}_1 \rangle = c_1$

$\langle$  $, \mathbf{h}_2 \rangle = c_2$

$\langle$  $, \mathbf{h}_3 \rangle = c_3$

$\langle$  $, \mathbf{h}_4 \rangle = c_4$

$\langle$  $, \mathbf{h}_5 \rangle = c_5$

$\langle$  $, \mathbf{h}_6 \rangle = c_6$

$\langle$  $, \mathbf{h}_7 \rangle = c_7$

$\langle$  $, \mathbf{h}_8 \rangle = c_8$

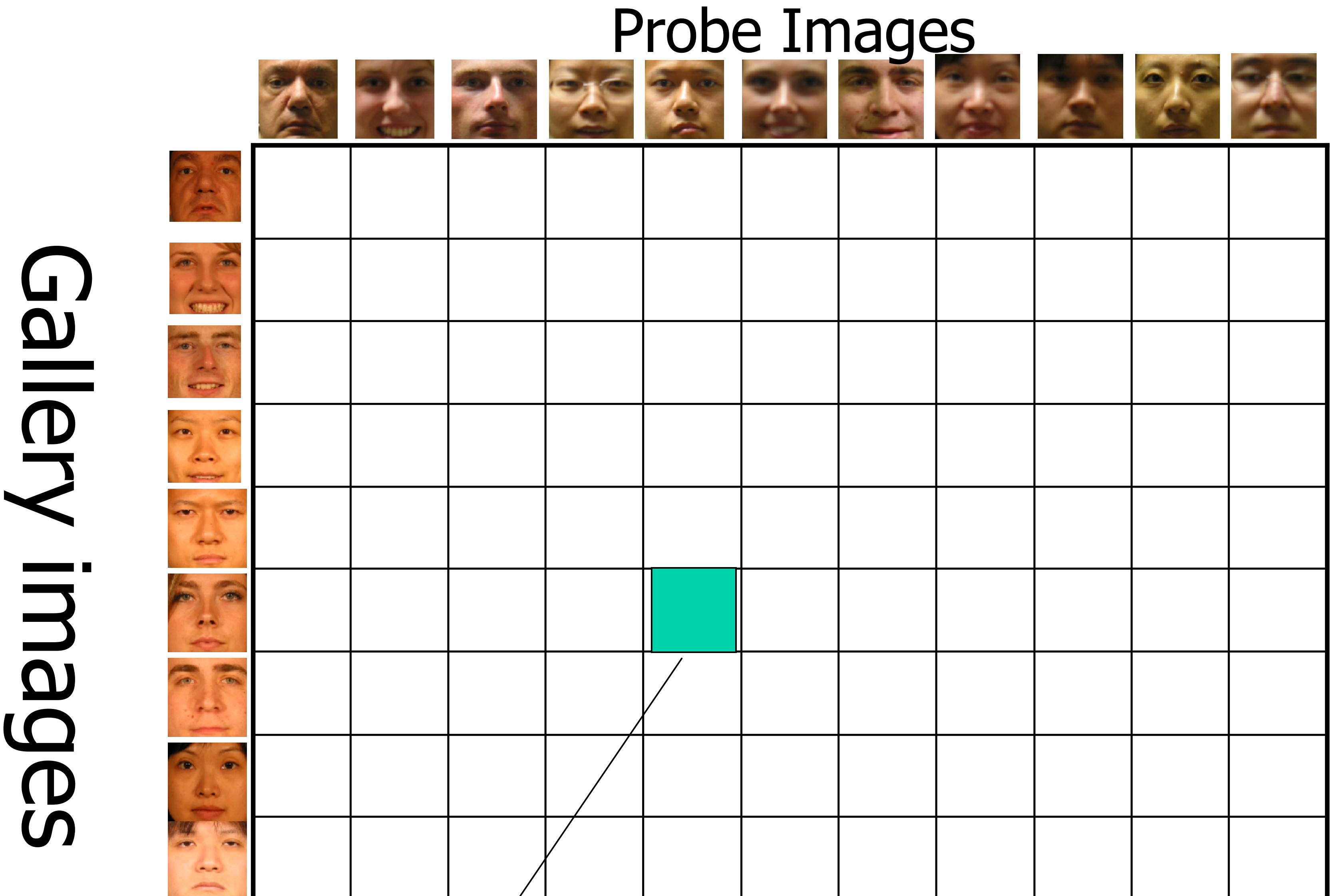
$\langle$  $, \mathbf{h}_N \rangle = c_N$

$$\left.\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \\ c_5 \\ c_6 \\ c_7 \\ c_8 \\ c_N \end{pmatrix}\right\} \mathbf{c}$$

Filter Design 1…222

- We design 222 filters based on the 222 generic subjects in the generic dataset.
- Each person has variable number of images.
- Correlation plane energy is minimized for all images in the MACE filter design.
- Much more efficient dimensionality reduction compared to PCA which needs to keep over 2000 eigenvectors.

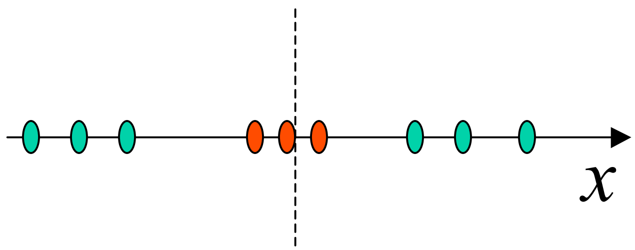
Populating Similarity Matrix



Similarity between KCFA() and KCFA() cosine distance is used

Nonlinear Feature Mapping Representations (CFs -> Kernel Cfs)

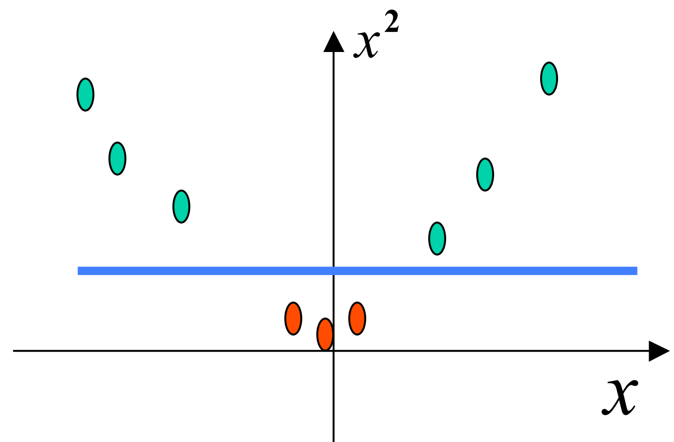
Linearly not separable problem may be separable in higher dimensional spaces.



Not Linearly Separable

$$\Phi : R^1 \rightarrow R^2$$

$$(x) \rightarrow (x, x^2)$$



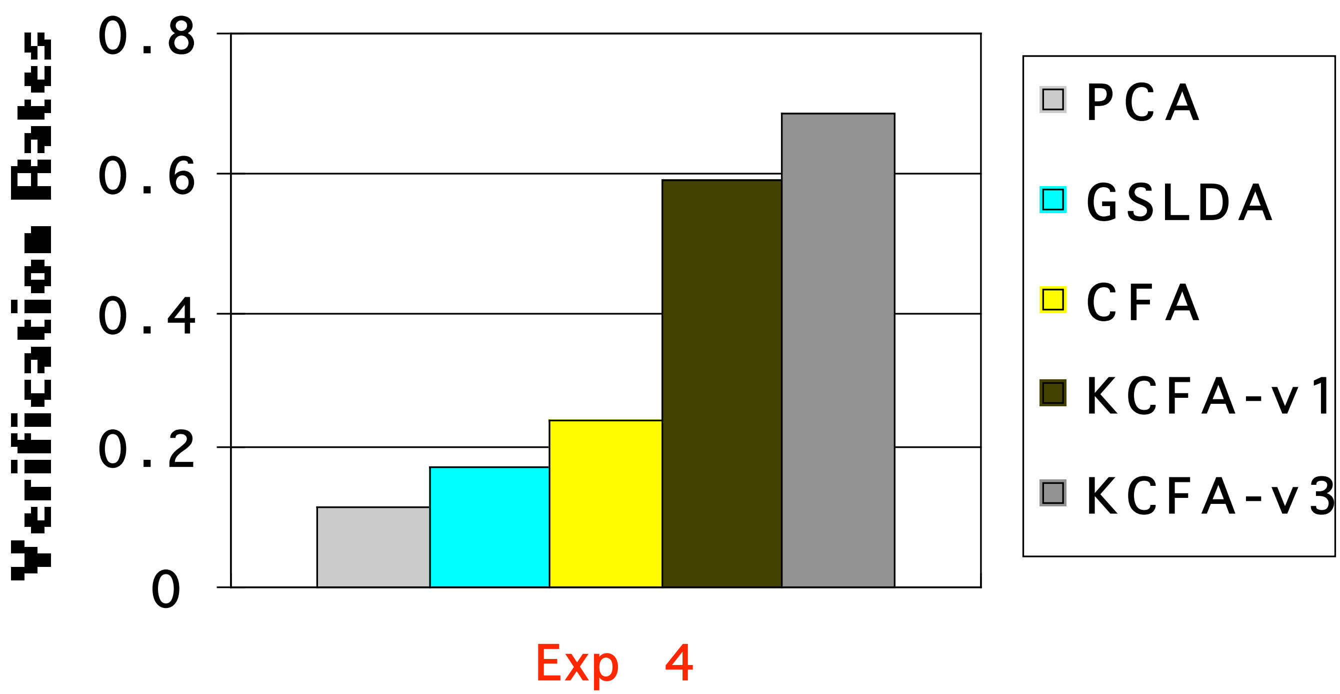
Linearly Separable

It will be intractable to compute the high dimensional space
-> Kernel tricks enable it without computing actual mapping

- Kernels form a dot (inner) product $K(x,y) = <\Phi(x), \Phi(y)>$

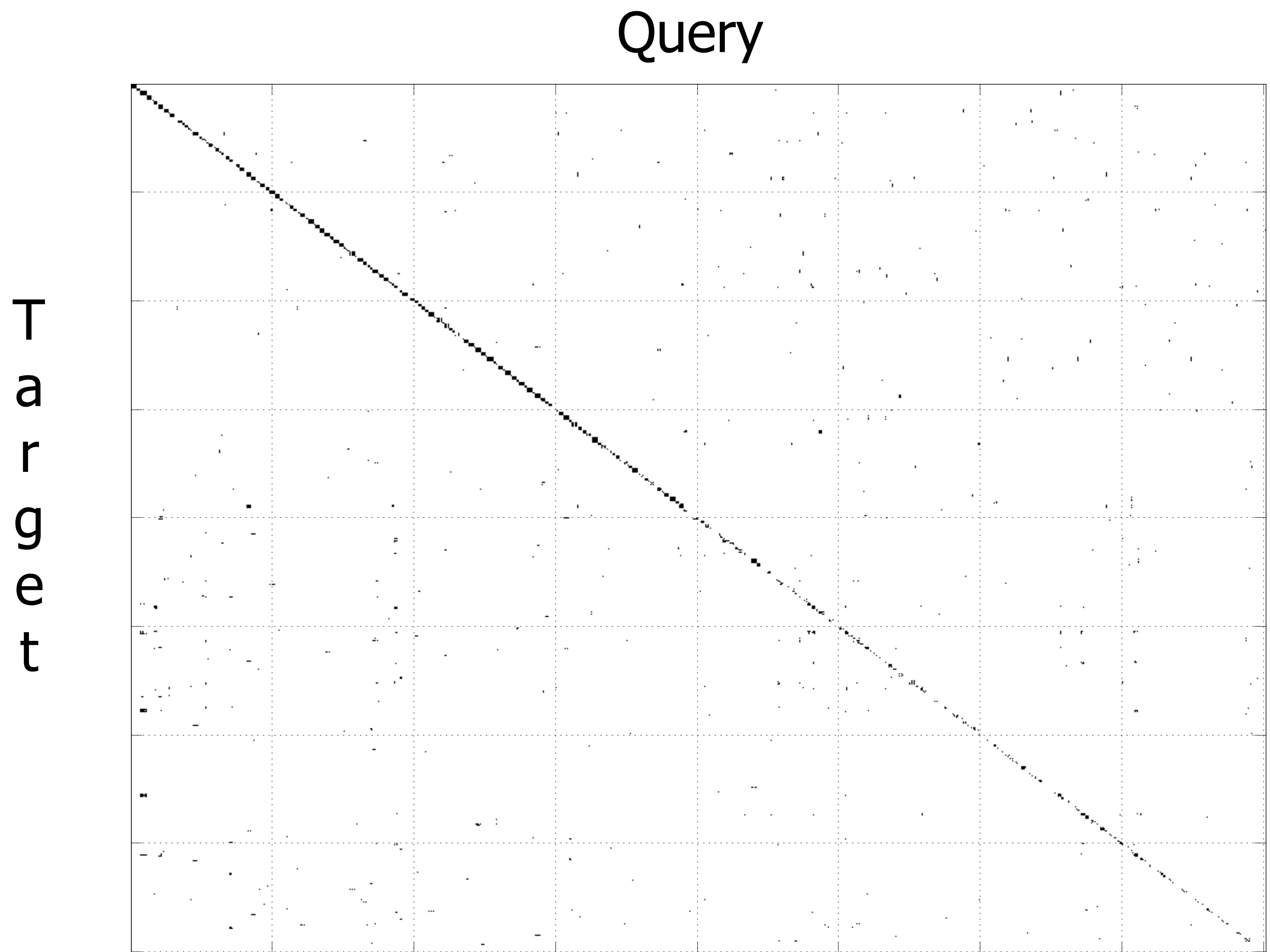
Experimental Results

- Eigenfaces (Baseline) results are provided by FRGC teams
- Performance measured at 0.1 % FAR (False Acceptance Rate)

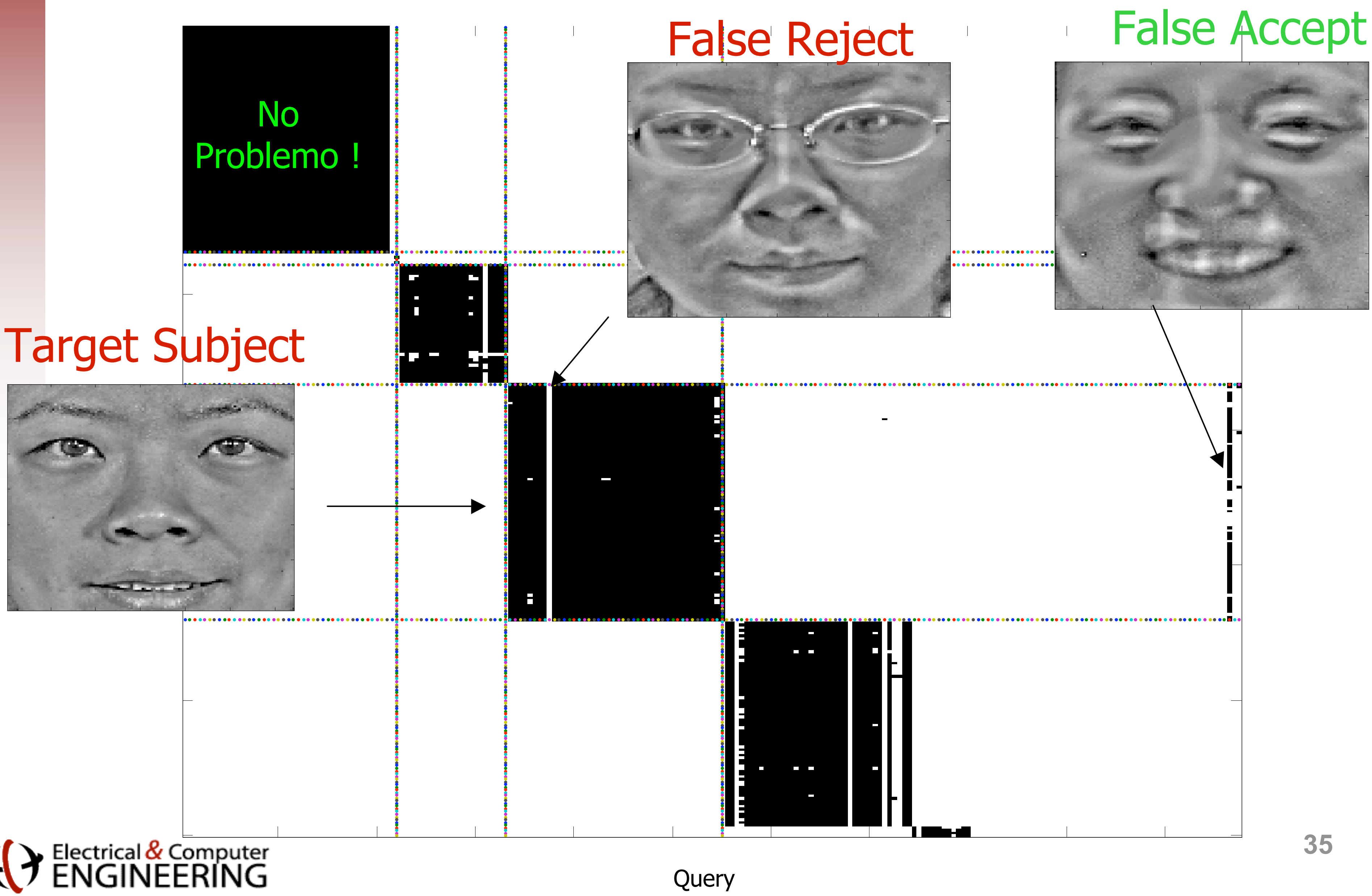


Similarity Matrix (of our poor performance algo)

(cosine distance-> thresholded 68%VR @ 0.1% FAR)



First Four classes



A closer look....

TARGET

#4202

Well classified PROBE

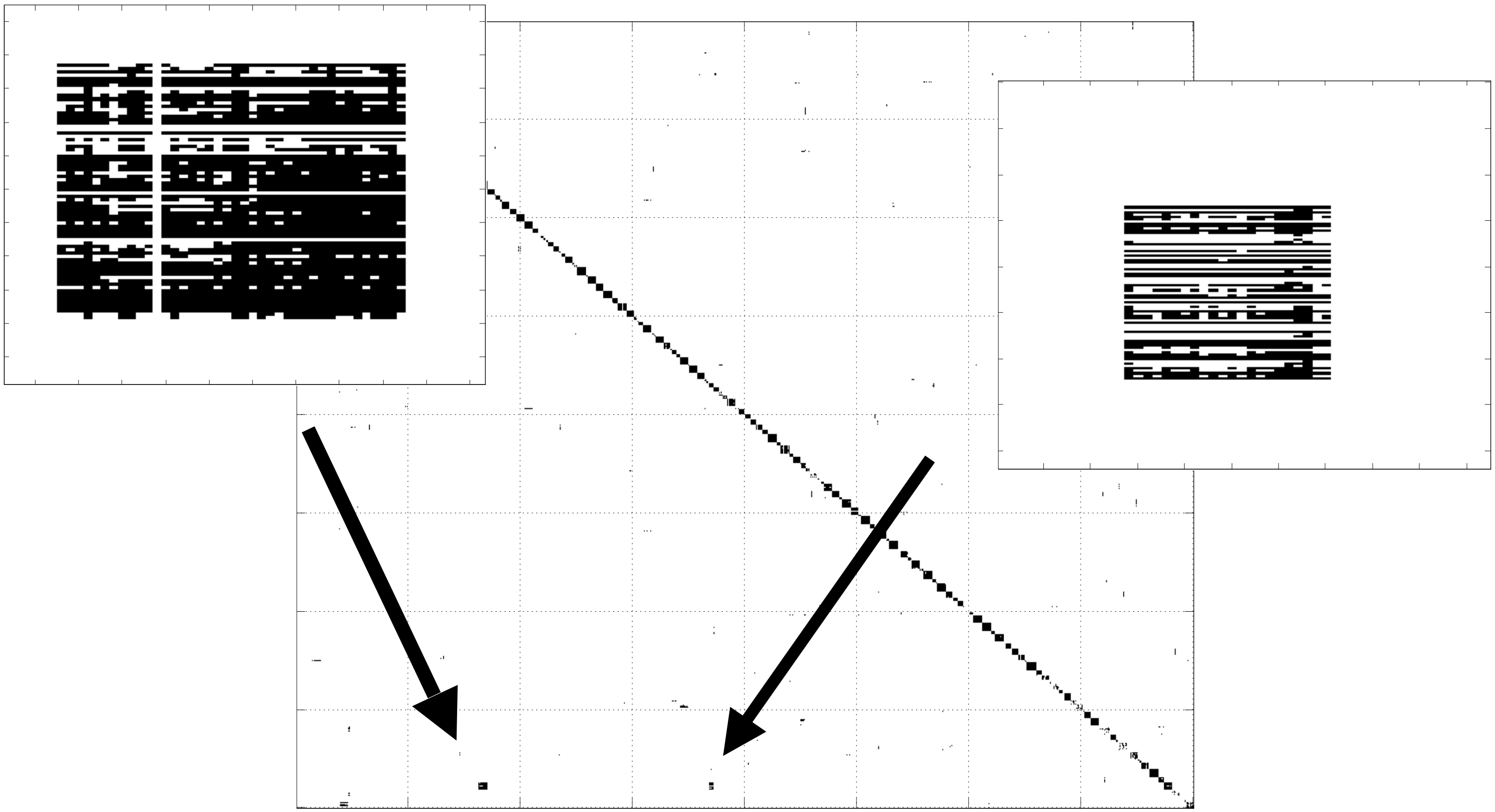
False Reject (some rotation, scale, due inaccurate eye-location

Well classified PROBE

False Accept

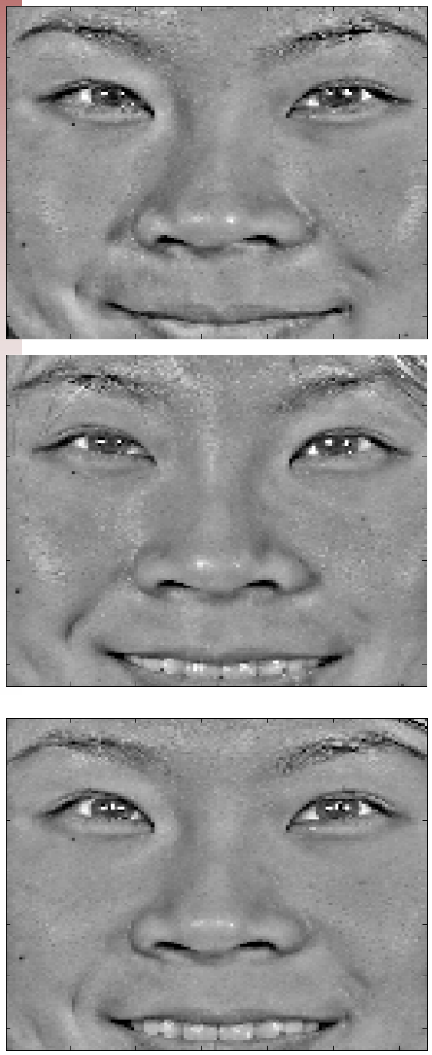
04221d452 impostor – Out of focus blur+harsh overhead illumination

False Accept Blocks (off-diagonal)



FA Block 1: they do look alike

Target Subject



False Accepts

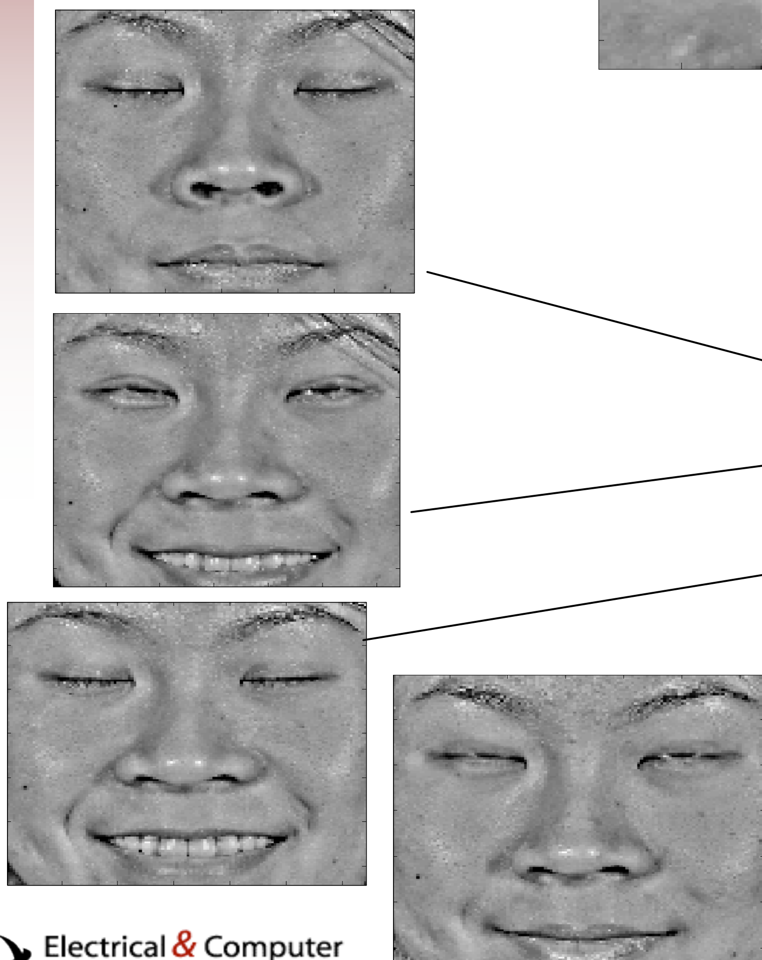
4629 not in generic

4334 in training

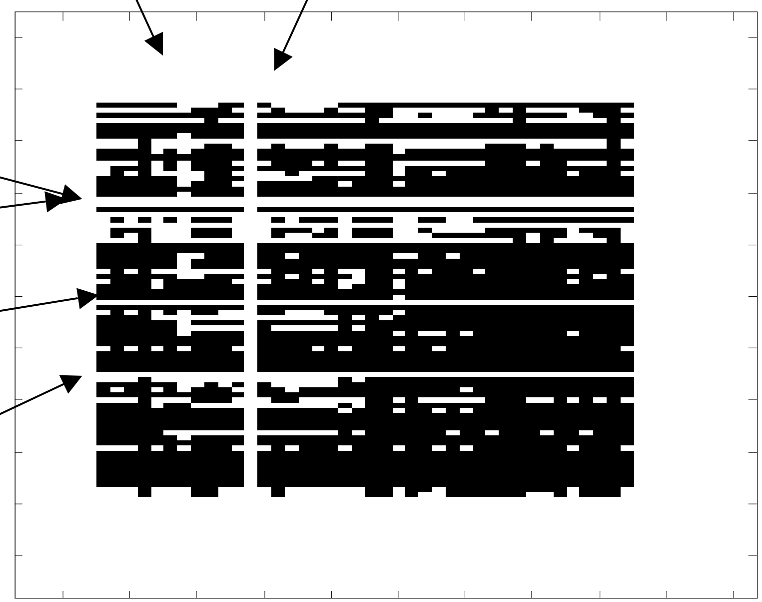
WEAK IMPOSTORS

TARGET :

Eyes closed



False Accept Block



Closing her eyes actually stopped False Accepts!

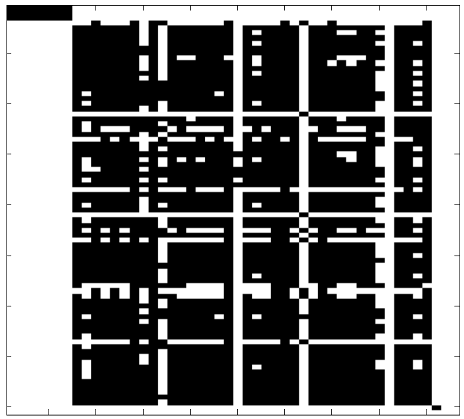
FA Block 2: same culprit

Target Subject

False Accepts (features match up)…

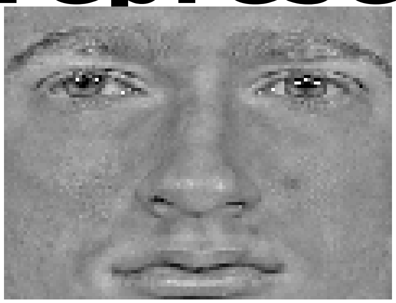
What is worrying part is….the False Accept image is male! So other domain knowledge (gender classification) can be helpful in such cases.

Target Subject however does a decent job in matching with her Query images



She does pretty good

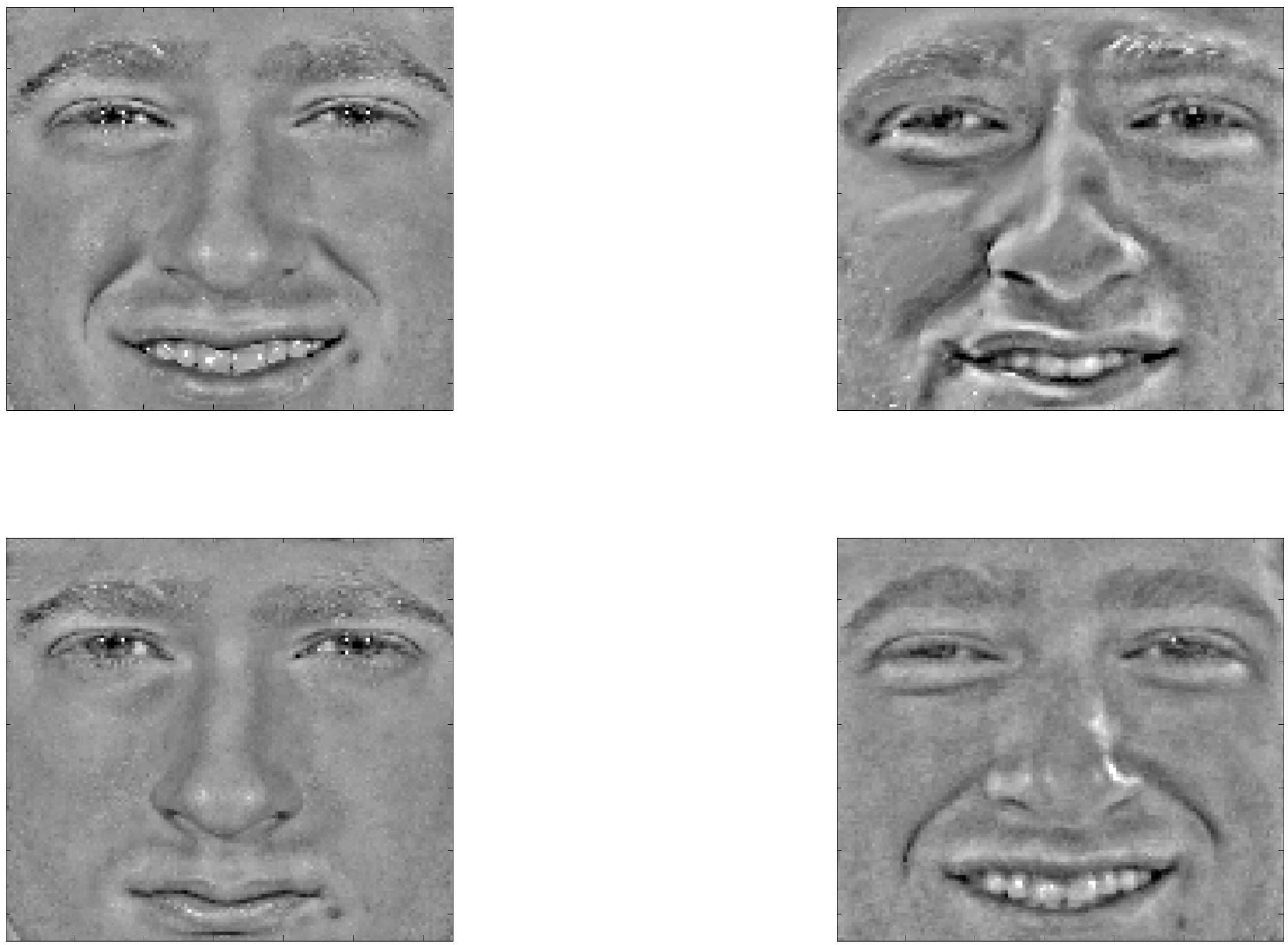
Mismatch Case (not well represented in training set?)



Target set

Query Set

These actually matched



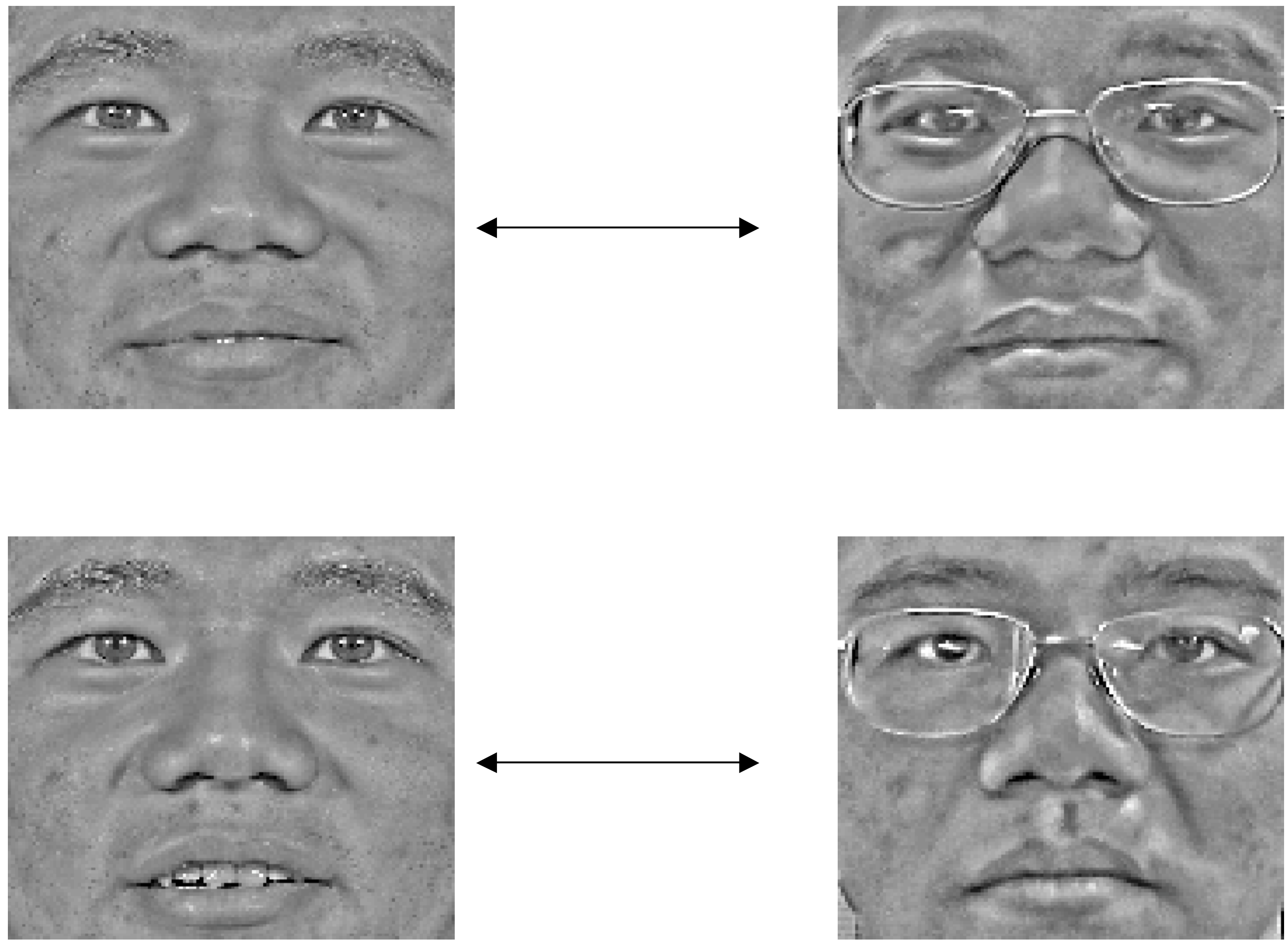
Problem case

4719 not in generic training set, is this why?

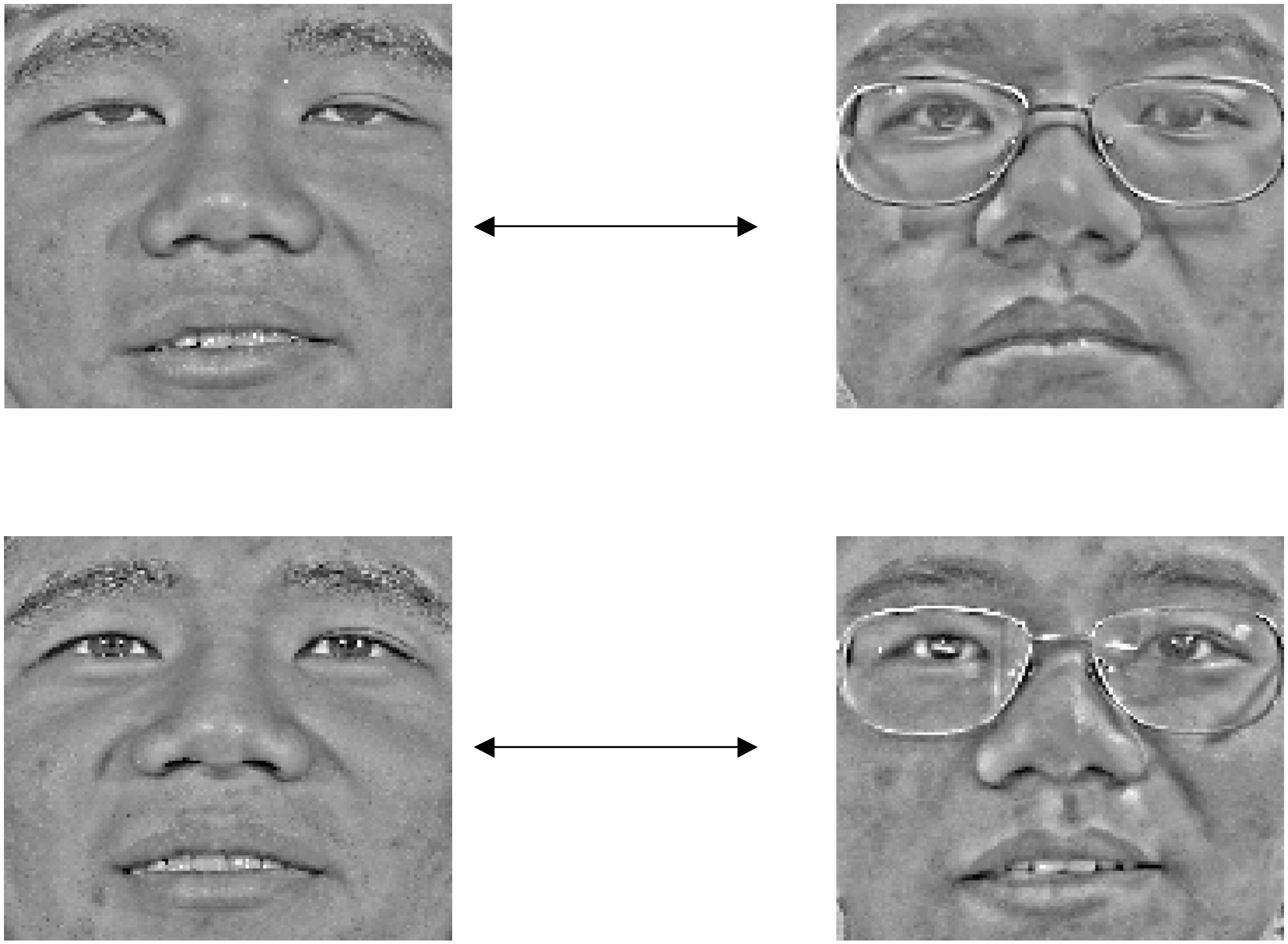
Another Target with matching problems

Corresponding Query images

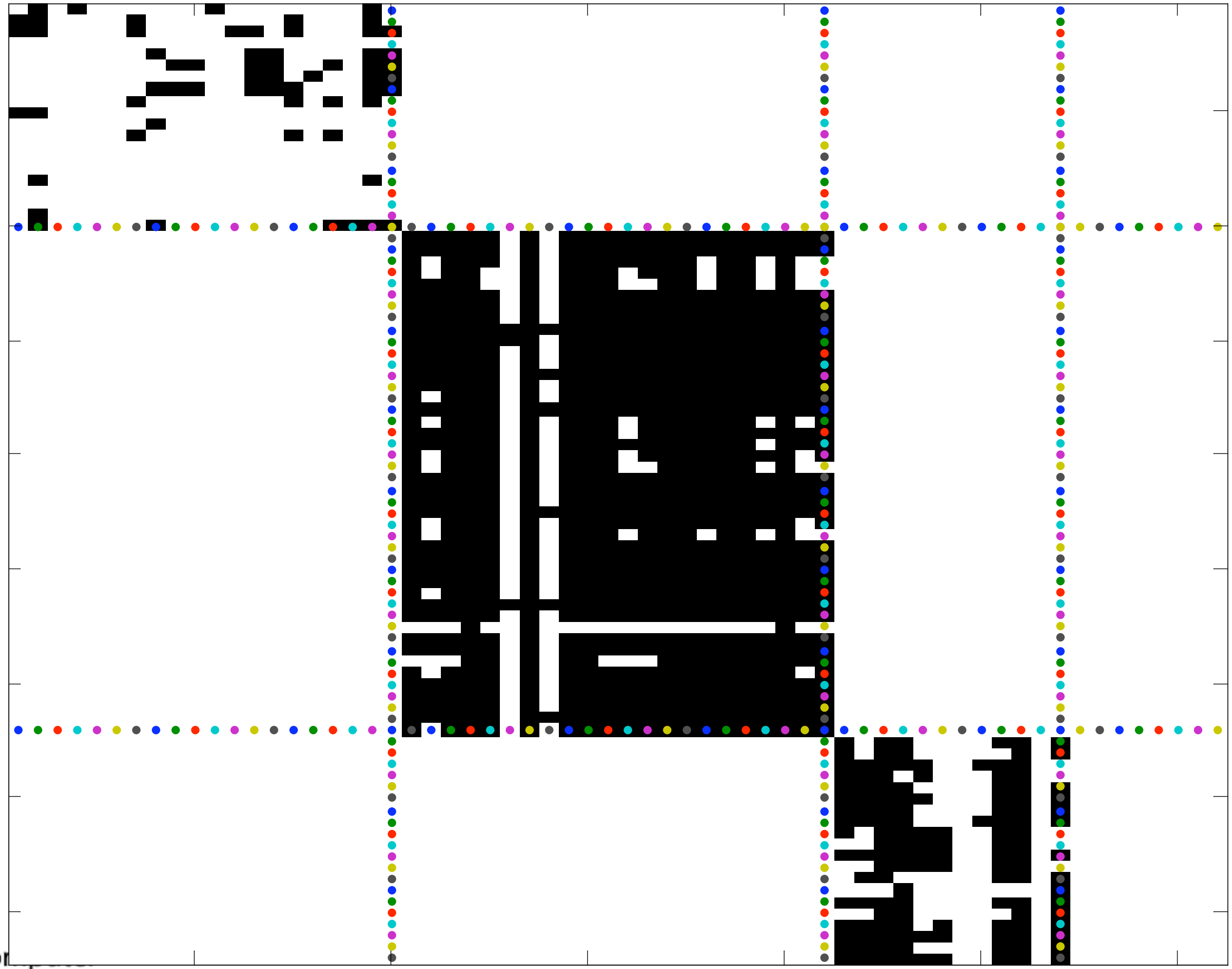
Some Examples of successful match (even with eye-glasses present)



Another few

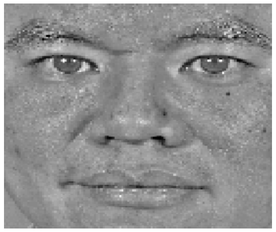


How about a person that was not modelled in generic and did great...?

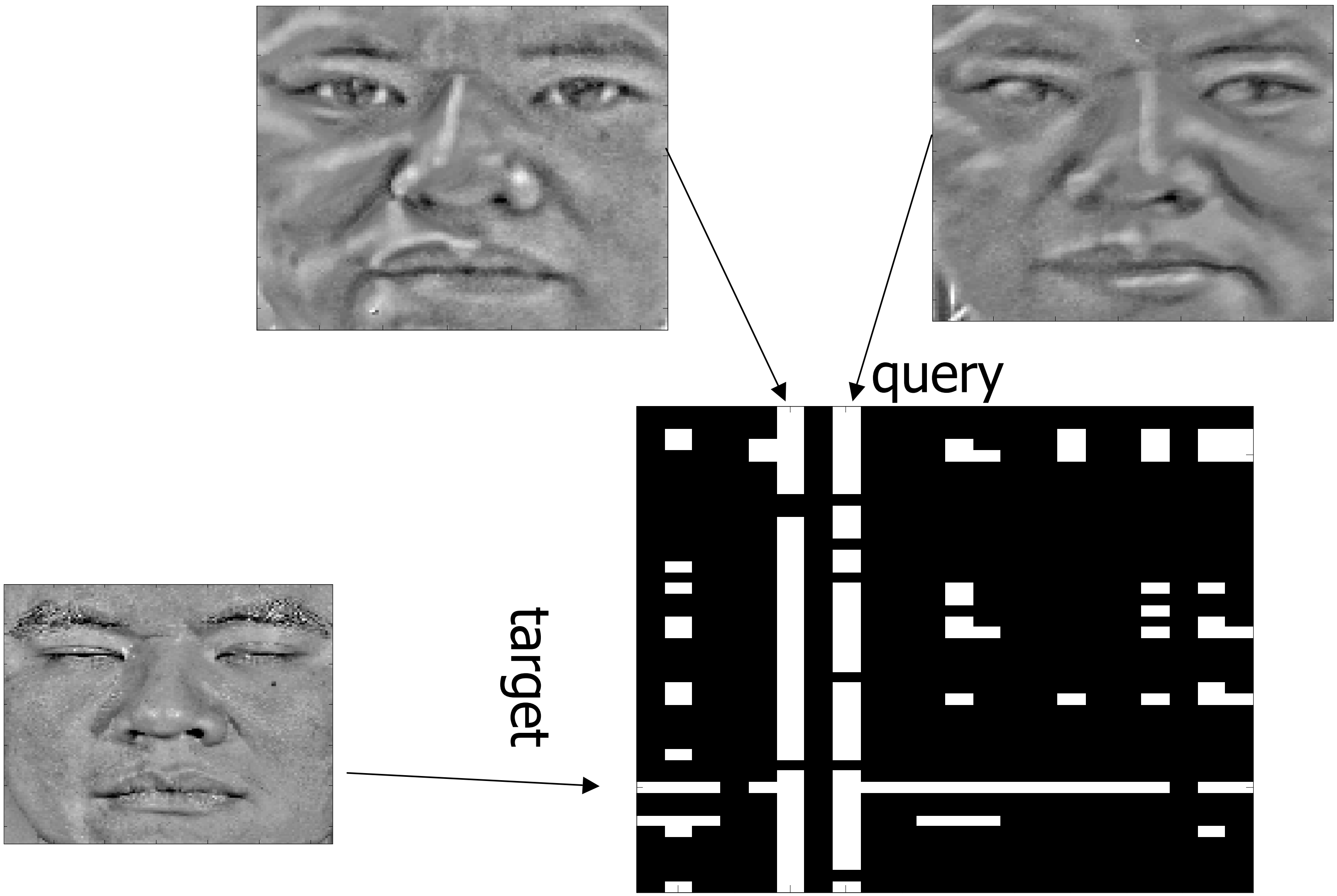


Person# 4866 Query that always matched

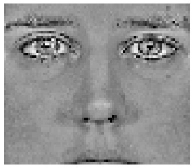
His target set



The ones that didn't match



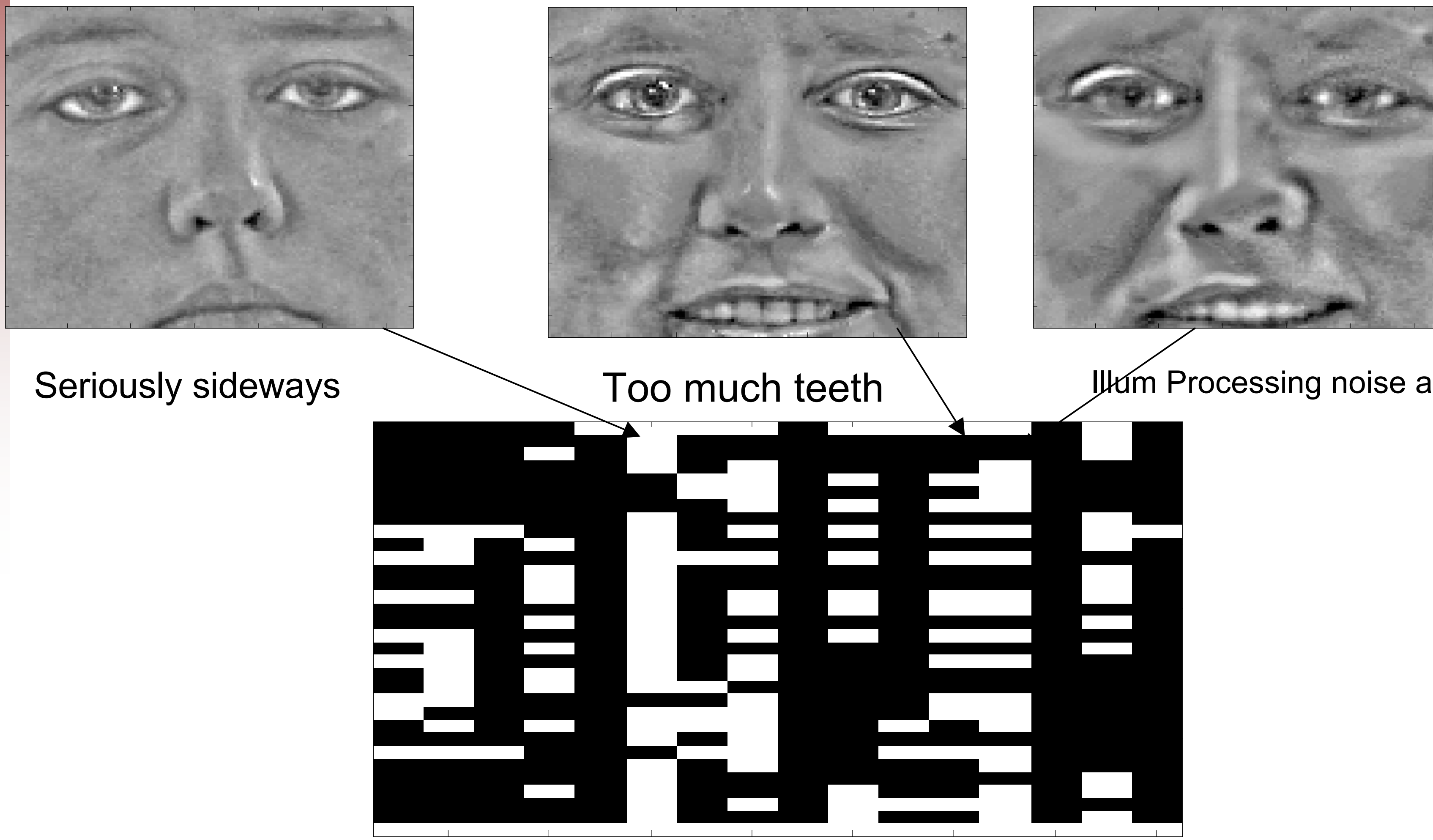
Another great match



Same cut lip crop issue

Random probe

Where it went wrong

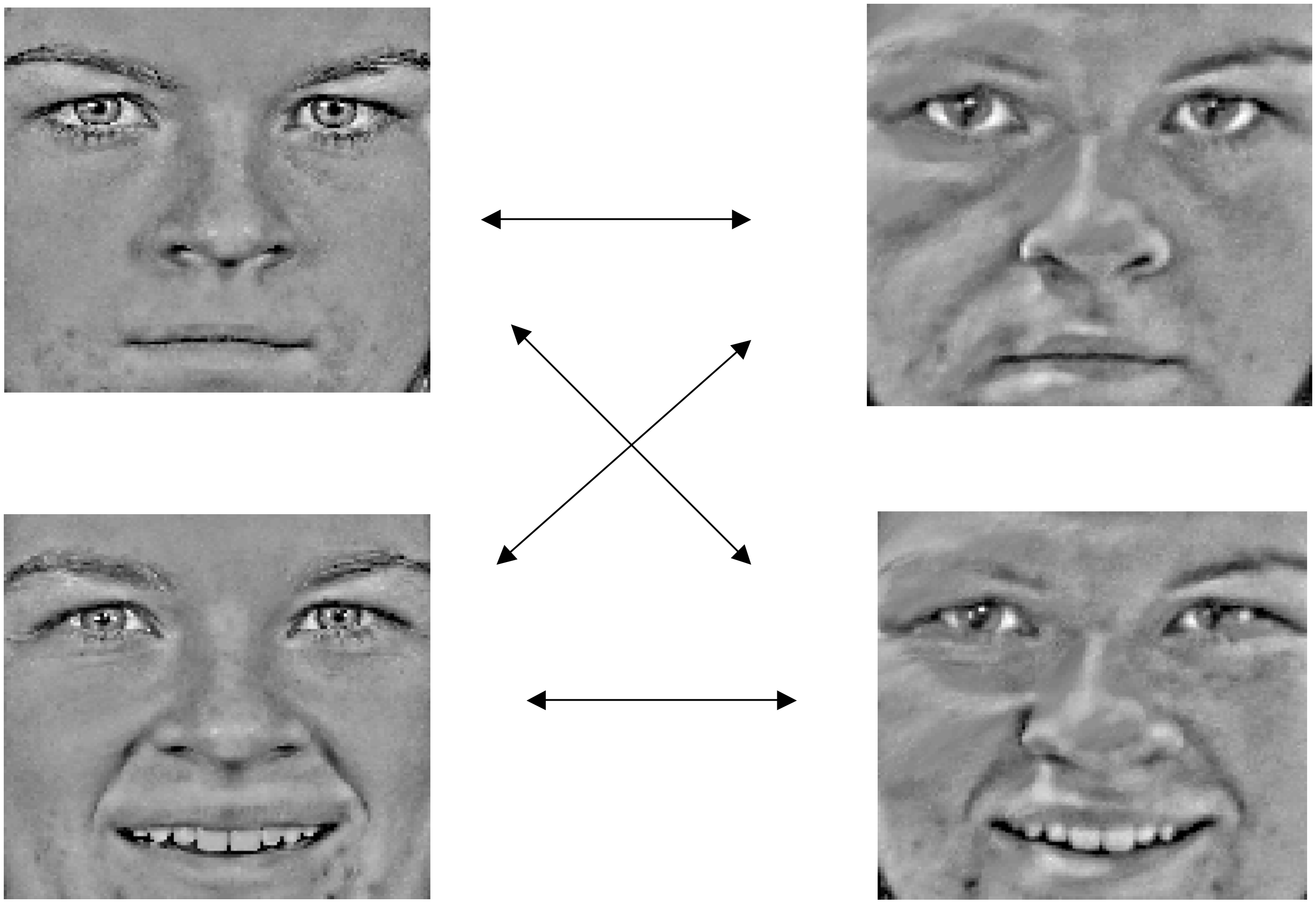


Another Random target class

Some Queries that did not match

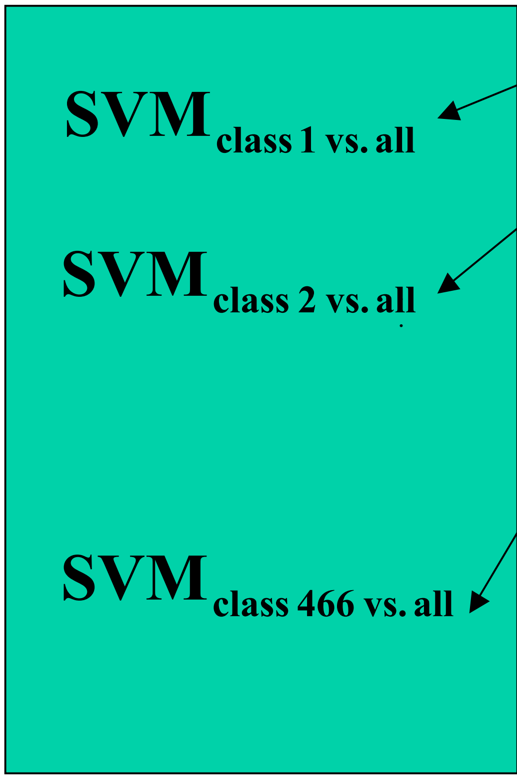
Never matched

Where it matched



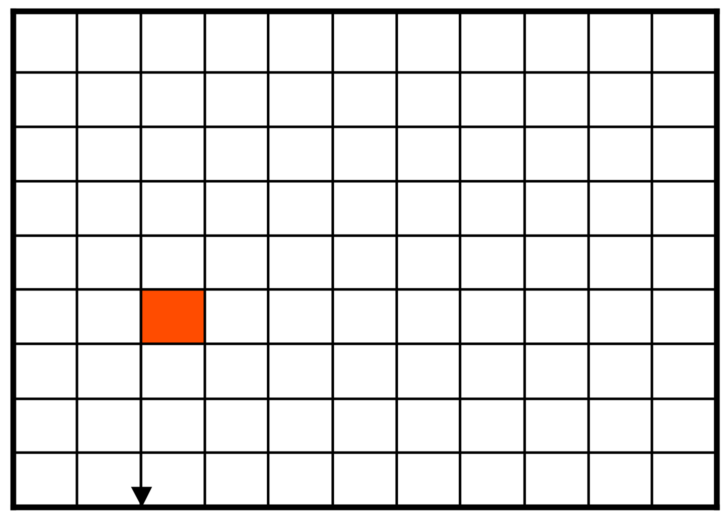
Distance Measure using Support Vector Machines

Gallery – 466 subjects (16,028 controlled images)



Probe – 466 subjects (8,014 uncontrolled images)

Similarity Matrix (16.028 X 8.014)

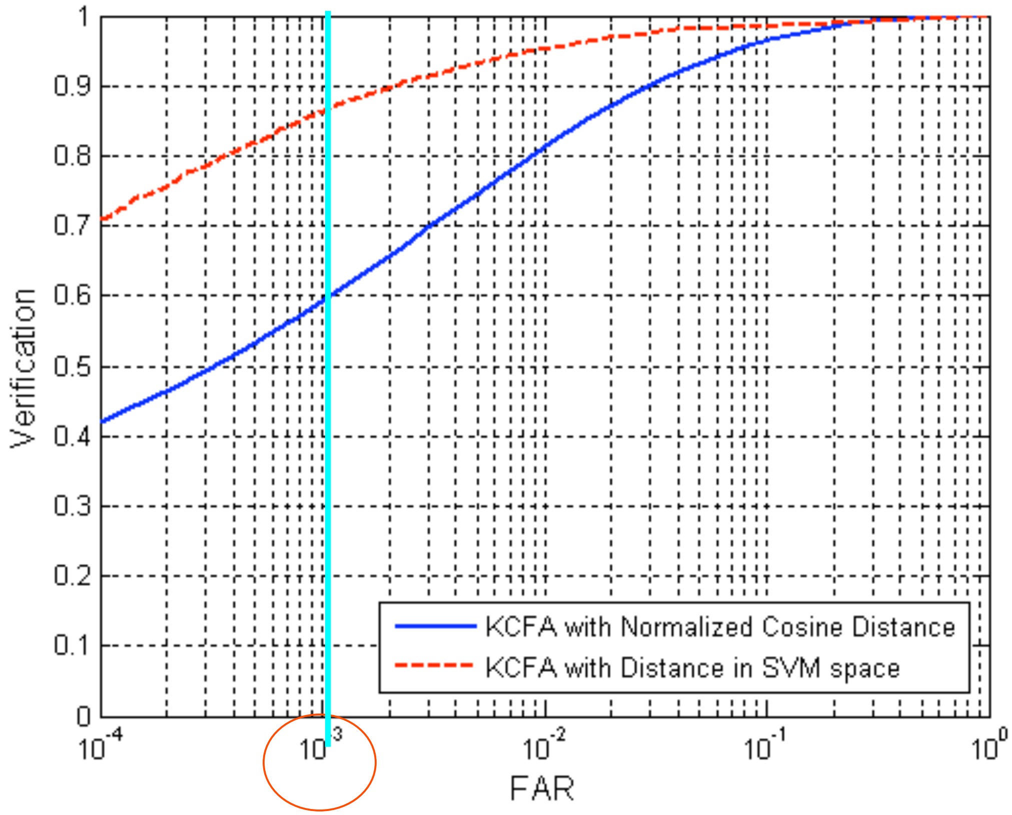


Distance between

6 (gallery) vs. 3(probe)

New Results

- **Using distance measure using the SVM improves the performance (KCFAv1 shown below)**



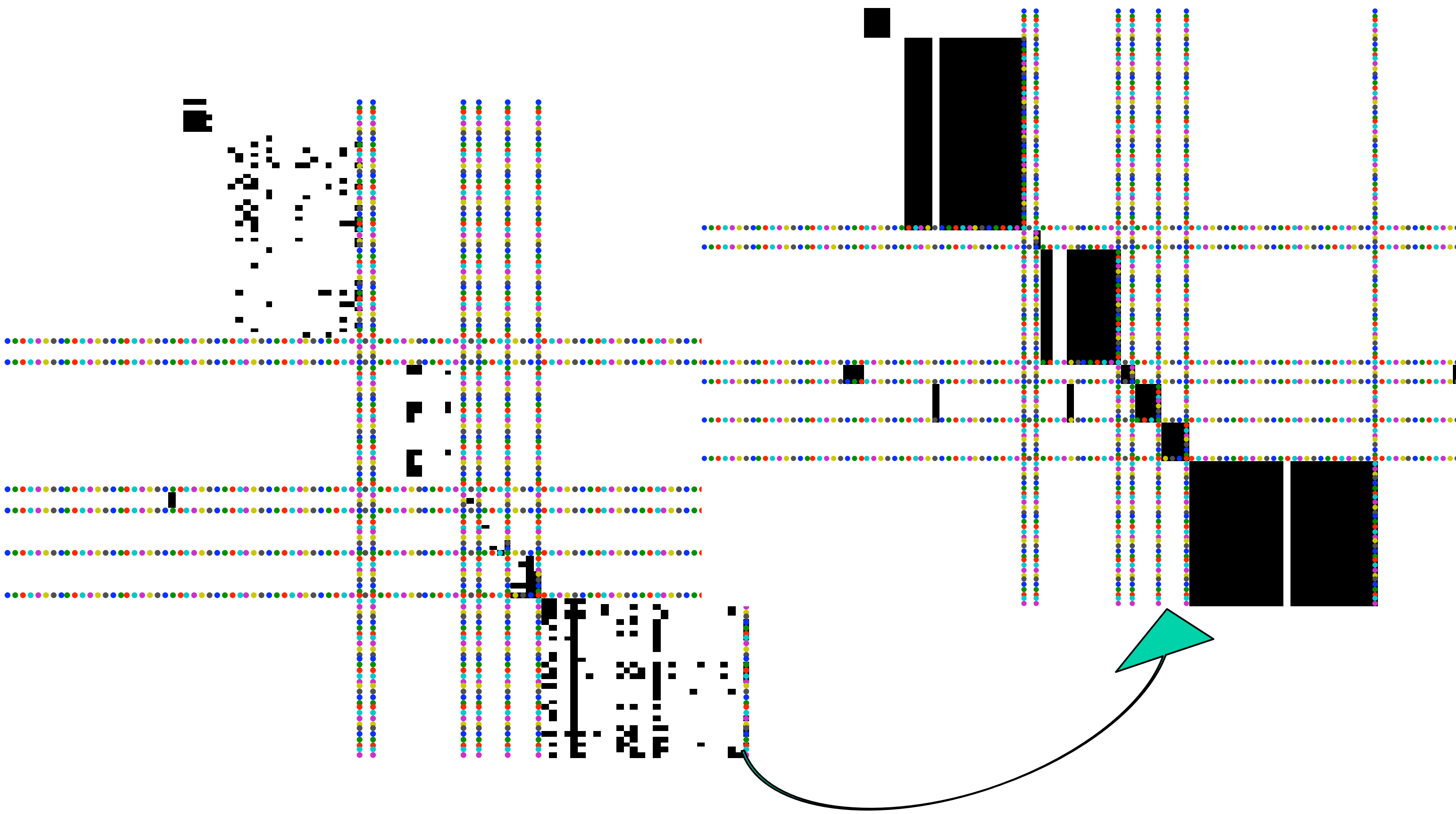
FAR at 0.1%

Electrical & Computer ENGINEERING

KCFA-SVM really improved performance

BEFORE: Part of similarity matrix using cosine distance only

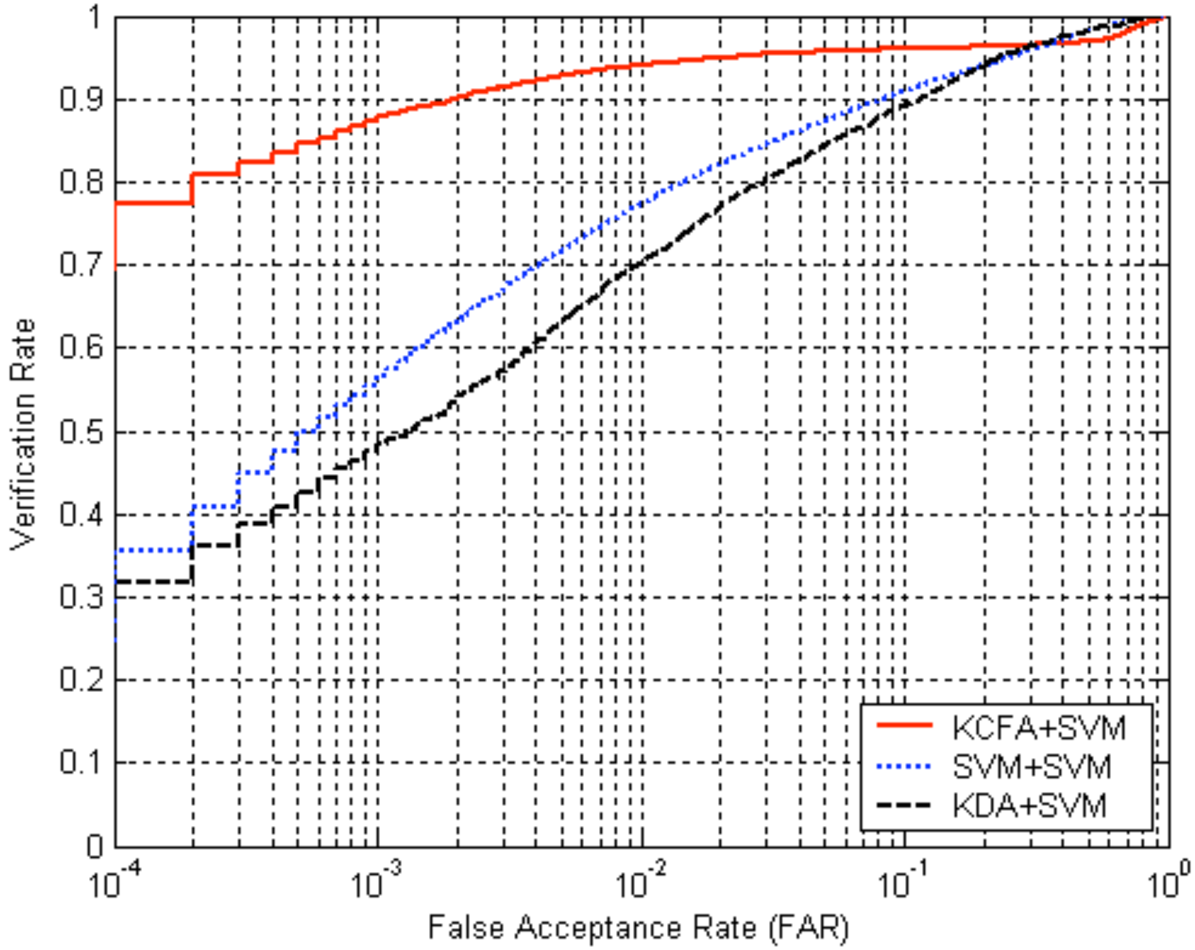
AFTER: Part of similarity matrix using KCFA+SVM distance



Other comparisons

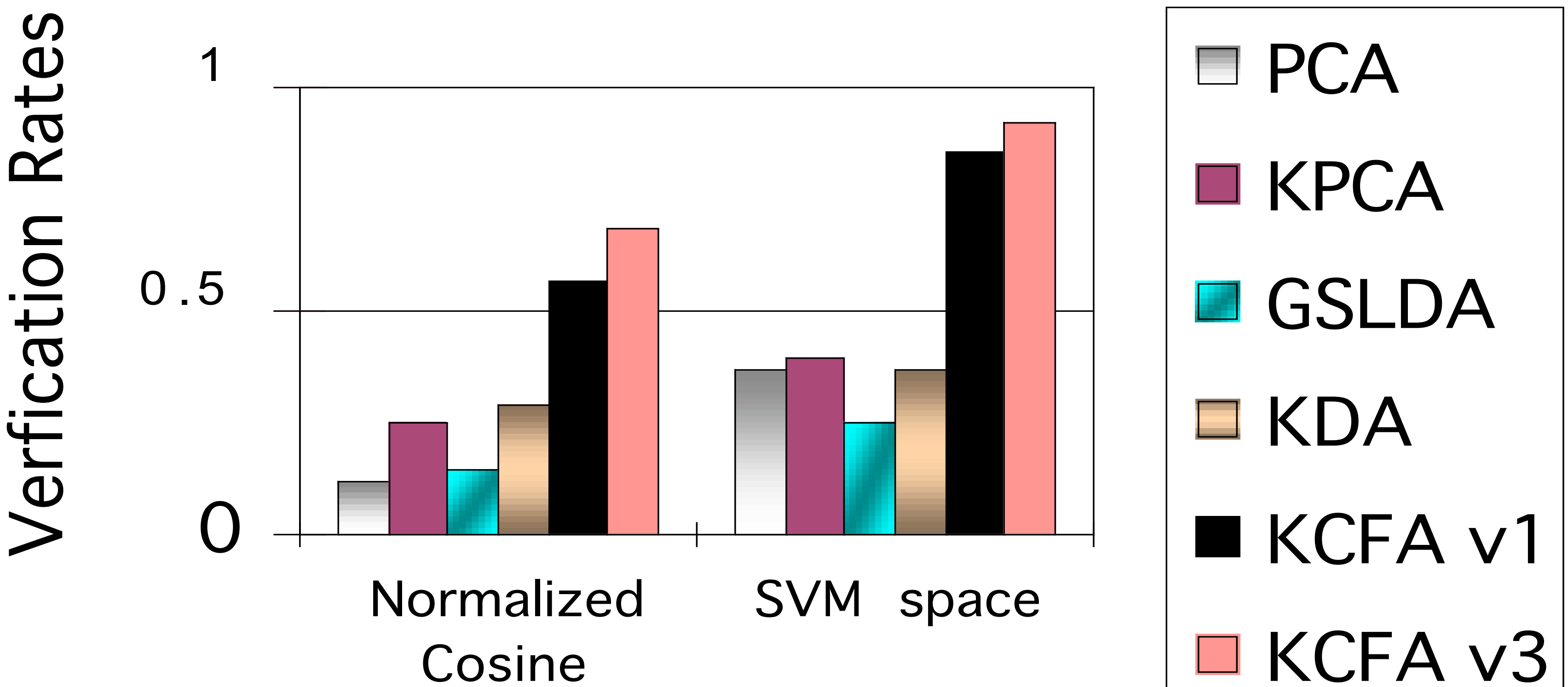
- What if we didn't use Correlation Filters in CFA framework?
- Are Correlation Filters the Optimal Classifiers to use?
- Use SVM in CFA discriminative framework and see what happens.
- Use Kernel Discriminant Analysis (KDA) in CFA discriminative framework

Different CFA Classifiers Benchmarks



Final Results

- **Using distance measure in the SVM space improve the performance regardless of the algorithms**



All results are based on using the exact same processed image data.

Which Facial Regions are most discriminative?

- Interesting question to pose…which facial parts contain most discrimination or consistency.
- Do such an analysis on FRGC is great chance to make some observations based on large amounts of face data
- We split FRGC data into 3 facial parts:
  - Eye Region
  - Nose Region
  - Mouth Region
- Do KCFA analysis on each region and analyze performance results.

Facial Regions used

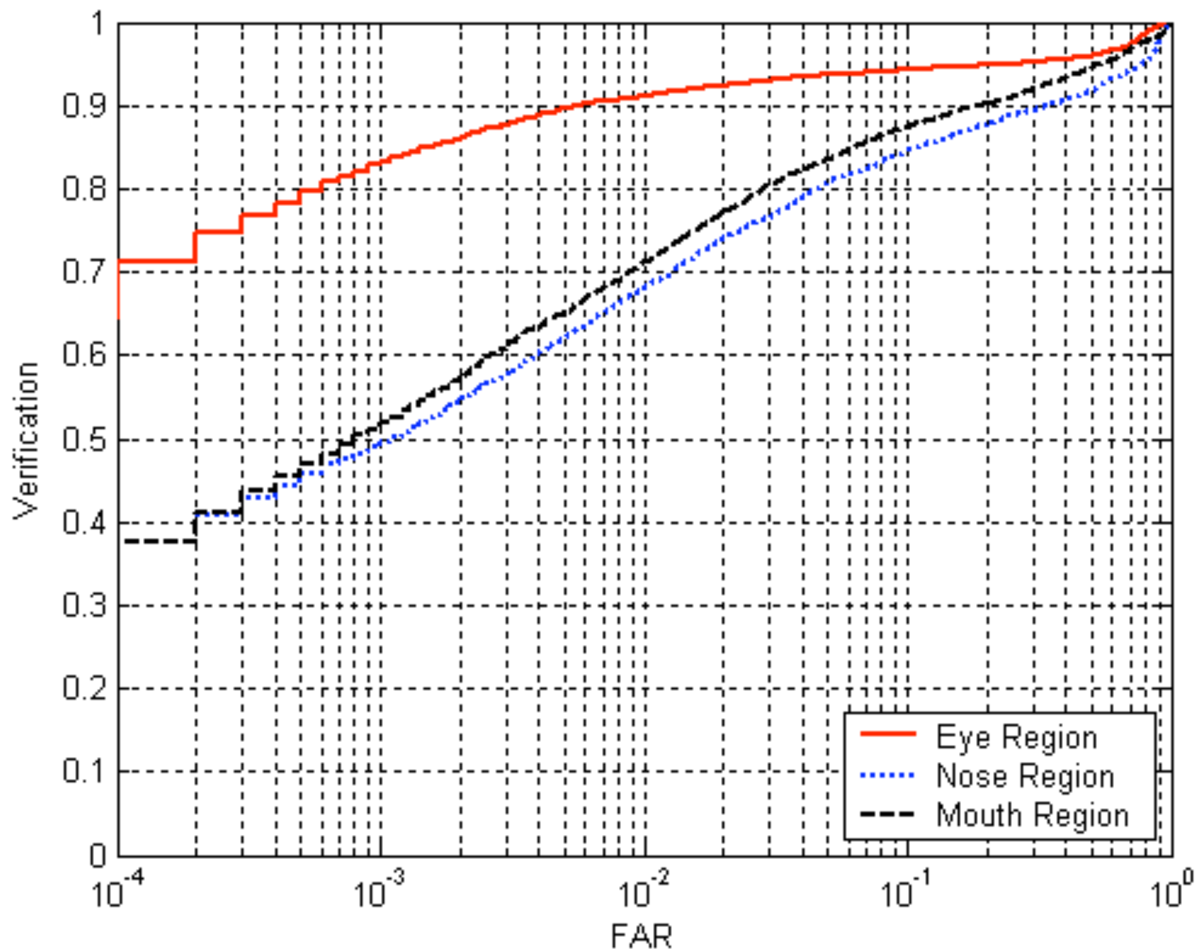
- Eye Region

- Nose Region

- Mouth Region

Experiments

- SVM-KCFA$_{v1}$ experiments for each face region.

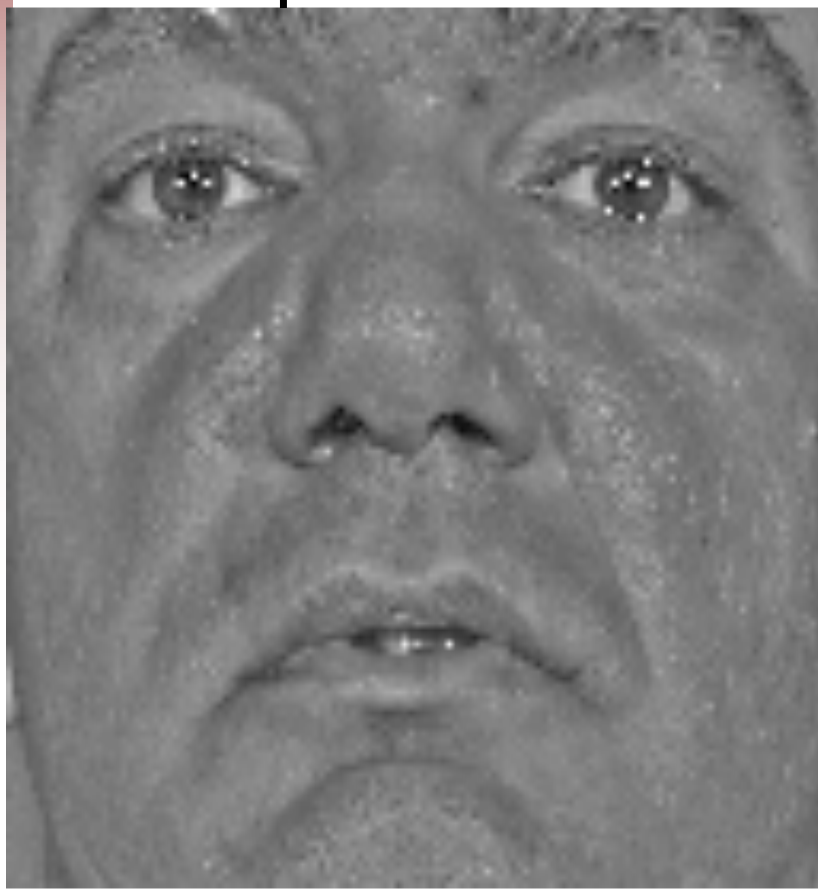


Discriminative Regions

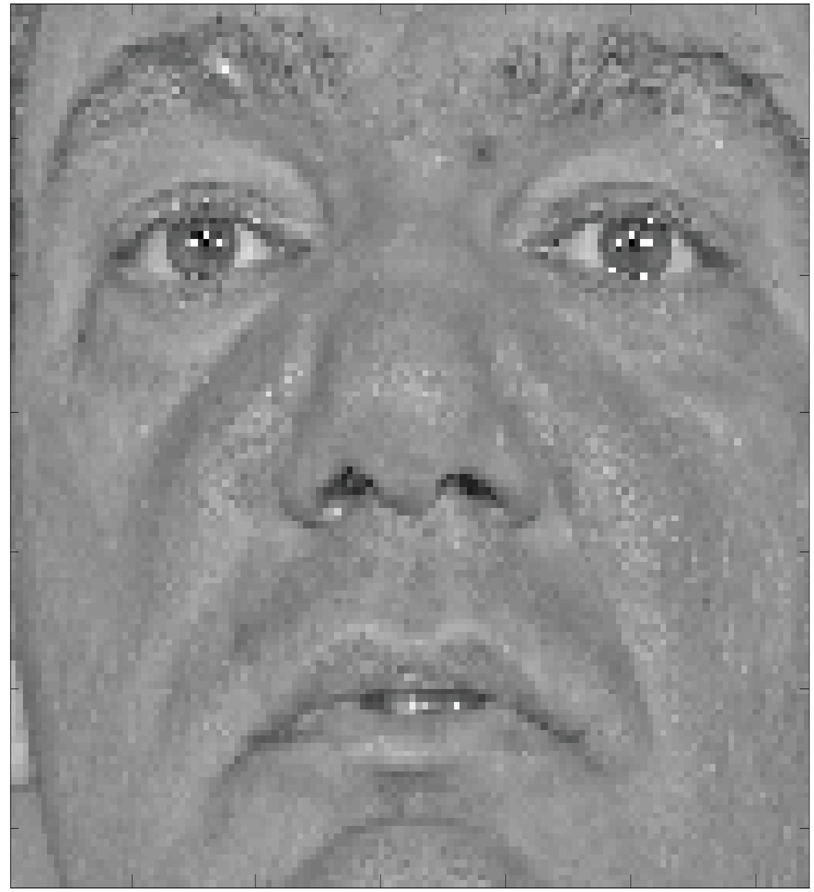
- We observe that eye-region is most discriminative (83.1%, 85% @ 0.1% FAR for $KCFA_{v1,}$ and $KCFA_{v3}$)
- Almost gives performance as using whole face region.
- Make intuitive sense as facial expressions change regions around the mouth and nose much more than eye-region.
- Fusion of $KCFA_{v1}$ features from eye-region and face-region increase performance to nearly ~90% @ 0.1 FAR. Now computing $KCFA_{v3}$ fusion to see performance boost.

What about registration errors?
(not using FFTs for shift-invariance)

3 pixels shift



10 pixels shift



Effect on Performance on FRGC2?

68%->64% @0.1 FAR (cos distance)

91.3%-> ~89% @ 0.1FAR (SVM)

Work to be done.

- Must improve illumination pre-processing. Use quality measure to weight scores.
- Pose estimation and correction, can be factored in image quality and recognition confidence.
- Quality can also be linked to :
  - Image alignment: as this affects matching performance
    - Robust eye-localization for scale and rotation normalization of face
    - Other fiducial points used for face registration.
    - Blur/out-of-focus
- Extract & use semantic information such as a gender classifier (and other information) to possibly reduce false matches. Mixture of experts (KCFA+?+?+?)
- Too many interesting & challenging things to do!...isn't this research field GREAT?!

Conclusion

- We can achieve ~92% using latest KCFA feature extraction methods (on whole face).
- Eye-region most discriminative, yields 85.1% alone! (working on getting new fusion results!)
- Substantial improvement over PCA which yields 12% @ 0.1% FAR
- Advanced Correlation Filters show superior performance yielding the best results as a CFA feature extraction classifier compared to other methods such as LDA, KDA, SVM.
- KCFA only extracted 222 features on all experiments, thus testing and discriminative learning is very fast and efficient!
- Our goal was to build the best “Core” matcher, now we can build more modules around this core matcher to improve performance:

We are very grateful to the United States Technical Support Working Group (TSWG) for funding and supporting this research effort

Special Thanks to:
Dr. David Herrington and Dr. Jim Wayman

Questions?